

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4670049号
(P4670049)

(45) 発行日 平成23年4月13日(2011.4.13)

(24) 登録日 平成23年1月28日(2011.1.28)

(51) Int. Cl. F I
G06F 13/00 (2006.01) G O 6 F 13/00 6 1 0 Q
H04L 12/58 (2006.01) H O 4 L 12/58 1 0 0 F

請求項の数 18 (全 40 頁)

(21) 出願番号	特願2005-150811 (P2005-150811)	(73) 特許権者	304027349
(22) 出願日	平成17年5月24日 (2005.5.24)		国立大学法人豊橋技術科学大学
(65) 公開番号	特開2006-260515 (P2006-260515A)		愛知県豊橋市天伯町雲雀ヶ丘1-1
(43) 公開日	平成18年9月28日 (2006.9.28)	(74) 代理人	110000534
審査請求日	平成20年5月23日 (2008.5.23)		特許業務法人しんめいセンチュリー
(31) 優先権主張番号	特願2005-39351 (P2005-39351)	(74) 代理人	100103045
(32) 優先日	平成17年2月16日 (2005.2.16)		弁理士 兼子 直久
(33) 優先権主張国	日本国(JP)	(74) 代理人	100127605
			弁理士 伊藤 愛
特許法第30条第1項適用	2004年8月20日 社	(74) 代理人	100129447
団法人電子情報通信学会発行の「FIT2004 第3			弁理士 橋本 努
回 情報科学技術フォーラム プログラム、講演論文集		(72) 発明者	山崎 仁
(CD-ROM)、一般講演論文集 第4分冊」に発表			愛知県豊橋市天伯町雲雀ヶ丘1-1
			国立大学法人豊橋技
			術科学大学内
			最終頁に続く

(54) 【発明の名称】 電子メールフィルタリングプログラム、電子メールフィルタリング方法、電子メールフィルタリングシステム

(57) 【特許請求の範囲】

【請求項1】

送信元から送信された不特定多数の電子メールに含まれる迷惑メールのフィルタリングを制御装置に実行させるための電子メールフィルタリングプログラムにおいて、

送信元から送信された電子メールのヘッダ情報を参照して、その電子メールの配信経路上における少なくとも1の中継装置のアドレスを取得する中継装置アドレス取得ステップと、

前記配信経路上における中継ルータのアドレスを取得する中継ルータアドレス取得ステップと、

その中継ルータアドレス取得ステップにより得られた中継ルータのアドレスに基づいて、前記中継装置アドレス取得ステップにおいてアドレスが取得された中継装置以外の中継装置であって前記配信経路を補完する中継装置のアドレスを取得する補完アドレス取得ステップと、

前記中継装置アドレス取得ステップにおいてアドレスが取得された中継装置と、前記補完アドレス取得ステップにおいてアドレスが取得された中継装置とに対し、それらの中継装置が過去に中継した迷惑メール及び正当なメールの各頻度を示す情報であって情報記憶手段に記憶されているメール情報に基づいて、該中継装置によって中継された電子メールが迷惑メールである確率を、ベイズ確率モデルを用いて得る迷惑メール中継確率取得ステップと、

その迷惑メール中継確率取得ステップにおいて得られた確率に基づいて、前記送信元か

10

20

ら送信された電子メールが迷惑メールである確率を得る迷惑メール受信確率取得ステップと、

その迷惑メール受信確率取得ステップにおいて得られた確率に応じて前記電子メールを所定区分に分類するメール判定ステップとを備えていることを特徴とする電子メールフィルタリングプログラム。

【請求項 2】

前記メール判定ステップは、

前記迷惑メール受信確率取得ステップにおいて得られた確率が、第 1 閾値を越えたか又は第 1 閾値以上であった場合に、前記送信元から送信された電子メールを迷惑メールであると判定する迷惑メール判定ステップと、

10

前記迷惑メール受信確率取得ステップにおいて得られた確率が、第 2 閾値未満又は第 2 閾値以下の場合に、前記送信元から送信された電子メールを正当なメールであると判定する正当メール判定ステップとを含み、

前記迷惑メール判定ステップ又は前記正当メール判定ステップにおいて判定された結果に応じて、前記中継装置アドレス取得ステップ又は補完アドレス取得ステップにおいて取得されたアドレスの中継装置について、前記情報記憶手段に記憶されている前記メール情報を更新する情報更新ステップを備えていることを特徴とする請求項 1 記載の電子メールフィルタリングプログラム。

【請求項 3】

前記メール判定ステップは、前記第 1 閾値と前記第 2 閾値とが異なる場合に、前記迷惑メール判定ステップ又は前記正当メール判定ステップにおいて前記迷惑メール又は前記正当なメールのいずれにも非該当であると判定された前記電子メールを、不確定メールと認識する不確定メール認識ステップをさらに含むことを特徴とする請求項 2 記載の電子メールフィルタリングプログラム。

20

【請求項 4】

前記メール判定ステップは、前記認識ステップにおいて不確定メールと認識された前記電子メールに対し、その電子メールに含まれるテキスト情報を利用することによって、その電子メールが迷惑メールであるか又は正当なメールであるかを判定する不確定メール再判定ステップをさらに含むことを特徴とする請求項 3 記載の電子メールフィルタリングプログラム。

30

【請求項 5】

前記中継装置アドレス取得ステップ又は補完アドレス取得ステップにより取得された中継装置のアドレスが正当なアドレスであることを確認するアドレス確認ステップを備えていると共に、

そのアドレス確認ステップにより前記アドレスが不正なアドレスであると確認された場合には、前記メール判定ステップにおいて、前記送信元から送信された電子メールが迷惑メールであると判定することを特徴とする請求項 2 から 4 のいずれかに記載の電子メールフィルタリングプログラム。

【請求項 6】

前記中継装置アドレス取得ステップ又は補完アドレス取得ステップにより取得されたアドレスの中継装置に対応する前記メール情報が前記情報記憶手段に不在の新出の中継装置である場合には、前記迷惑メール中継確率取得ステップにおいて、その新出の中継装置に対し、前記迷惑メールである確率として所定値を付与することを特徴とする請求項 1 から 5 のいずれかに記載の電子メールフィルタリングプログラム。

40

【請求項 7】

送信元から送信された不特定多数の電子メールに含まれる迷惑メールをフィルタリングすることができる電子メールフィルタリング方法において、

送信元から送信された電子メールのヘッダ情報を参照して、その電子メールの配信経路上における少なくとも 1 の中継装置のアドレスを取得する中継装置アドレス取得手段と、
前記配信経路上における中継ルータのアドレスを取得する中継ルータアドレス取得手段

50

と、

その中継ルータアドレス取得手段により得られた中継ルータのアドレスに基づいて、前記中継装置アドレス取得手段によりアドレスが取得された中継装置以外の中継装置であって前記配信経路を補完する中継装置のアドレスを取得する補完アドレス取得手段と、

前記中継装置アドレス取得手段によりアドレスが取得された中継装置と、前記補完アドレス取得手段によりアドレスが取得された中継装置とに対し、それらの中継装置が過去に中継した迷惑メール及び正当なメールの各頻度を示す情報であって情報記憶手段に記憶されているメール情報に基づいて、該中継装置によって中継された電子メールが迷惑メールである確率を、ベイズ確率モデルを用いて得る迷惑メール中継確率取得手段と、

その迷惑メール中継確率取得手段により得られた確率に基づいて、前記送信元から送信された電子メールが迷惑メールである確率を得る迷惑メール受信確率取得手段と、

その迷惑メール受信確率取得手段により得られた確率に応じて前記電子メールを所定区分に分類するメール判定手段とを備えていることを特徴とする電子メールフィルタリング方法。

【請求項 8】

前記メール判定手段は、

前記迷惑メール受信確率取得手段により得られた確率が、第 1 閾値を越えたか又は第 1 閾値以上であった場合に、前記送信元から送信された電子メールを迷惑メールであると判定する迷惑メール判定手段と、

前記迷惑メール受信確率取得手段により得られた確率が、第 2 閾値未満又は第 2 閾値以下の場合に、前記送信元から送信された電子メールを正当なメールであると判定する正当メール判定手段とを備えており、

前記迷惑メール判定手段又は前記正当メール判定手段により判定された結果に応じて、前記中継装置アドレス取得手段又は補完アドレス取得手段により取得されたアドレスの中継装置について、前記情報記憶手段に記憶されている前記メール情報を更新する情報更新手段を備えていることを特徴とする請求項 7 記載の電子メールフィルタリング方法。

【請求項 9】

前記メール判定手段は、前記第 1 閾値と前記第 2 閾値とが異なる場合に、前記迷惑メール判定手段又は前記正当メール判定手段により前記迷惑メール又は前記正当なメールのいずれにも非該当であると判定された前記電子メールを、不確定メールと認識する不確定メール認識手段をさらに備えていることを特徴とする請求項 8 記載の電子メールフィルタリング方法。

【請求項 10】

前記メール判定手段は、前記不確定メール認識手段により不確定メールと認識された前記電子メールに対し、その電子メールに含まれるテキスト情報を利用することによって、その電子メールが迷惑メールであるか又は正当なメールであるかを判定する不確定メール再判定手段をさらに備えていることを特徴とする請求項 9 記載の電子メールフィルタリング方法。

【請求項 11】

前記中継装置アドレス取得手段又は補完アドレス取得手段により取得された中継装置のアドレスが正当なアドレスであるかを確認するアドレス確認手段を備え、

そのアドレス確認手段により前記アドレスが不正なアドレスであると確認された場合に、前記迷惑メール判定手段は、前記送信元から送信された電子メールを迷惑メールであると判定することを特徴とする請求項 8 から 10 のいずれかに記載の電子メールフィルタリング方法。

【請求項 12】

前記迷惑メール中継確率取得手段は、前記中継装置アドレス取得手段又は補完アドレス取得手段により取得されたアドレスの中継装置に対応する前記メール情報が前記情報記憶手段に不在の新出の中継装置である場合には、その中継装置に対し、前記迷惑メールである確率として所定値を付与することを特徴とする請求項 7 から 11 のいずれかに記載の電

10

20

30

40

50

子メールフィルタリング方法。

【請求項 1 3】

電子メールを伝送可能な経路上において、送信元から送信された不特定多数の電子メールに含まれる迷惑メールをフィルタリングすることが可能な電子メールフィルタリングシステムにおいて、

送信元から送信された電子メールのヘッダ情報を参照して、その電子メールの配信経路上における少なくとも1の中継装置のアドレスを取得する中継装置アドレス取得手段と、

1の中継装置に対し、その中継装置によって過去に中継された迷惑メール及び正当なメールの各頻度を示すメール情報を記憶する情報記憶手段と、

前記配信経路上における中継ルータのアドレスを取得する中継ルータアドレス取得手段と、

その中継ルータアドレス取得手段により得られた中継ルータのアドレスに基づいて、前記中継装置アドレス取得手段においてアドレスが取得された中継装置以外の中継装置であって前記配信経路を補完する中継装置のアドレスを取得する補完アドレス取得手段と、

前記中継装置アドレス取得手段によりアドレスが取得された中継装置と、前記補完アドレス取得手段によりアドレスが取得された中継装置とに対し、情報記憶手段に記憶されている前記メール情報に基づいて、それらの中継装置によって中継された電子メールが迷惑メールである確率を、ベイズ確率モデルを用いて得る迷惑メール中継確率取得手段と、

その迷惑メール中継確率演算手段により得られた確率に基づいて、前記送信元から送信された電子メールが迷惑メールである確率を得る迷惑メール受信確率取得手段と、

その迷惑メール受信確率取得手段により得られた確率に応じて前記電子メールを所定区分に分類するメール判定手段とを備えていることを特徴とする電子メールフィルタリングシステム。

【請求項 1 4】

前記メール判定手段は、

前記迷惑メール受信確率取得手段により得られた確率が、第1閾値を越えたか又は第1閾値以上であった場合に、前記送信元から送信された電子メールを迷惑メールであると判定する迷惑メール判定手段と、

前記迷惑メール受信確率取得手段により得られた確率が、第2閾値未満又は第2閾値以下の場合に、前記送信元から送信された電子メールを正当なメールであると判定する正当メール判定手段とを備えており、

前記迷惑メール判定手段又は前記正当メール判定手段による判定結果に応じて、前記中継装置アドレス取得手段又は補完アドレス取得手段により取得されたアドレスの中継装置について、前記情報記憶手段に記憶されている前記メール情報を更新する情報更新手段を備えていることを特徴とする請求項1 3記載の電子メールフィルタリングシステム。

【請求項 1 5】

前記メール判定手段は、前記第1閾値と前記第2閾値とが異なる場合に、前記迷惑メール判定手段又は前記正当メール判定手段による判定が前記迷惑メール又は前記正当なメールのいずれにも非該当である電子メールを、不確定メールと認識する不確定メール認識手段をさらに備えていることを特徴とする請求項1 4記載の電子メールフィルタリングシステム。

【請求項 1 6】

前記メール判定手段は、前記不確定メール認識手段において不確定メールと認識された前記電子メールに対し、その電子メールに含まれるテキスト情報を利用することによって、その電子メールが迷惑メールであるか又は正当なメールであるかを判定する不確定メール再判定手段をさらに備えていることを特徴とする請求項1 5記載の電子メールフィルタリングシステム。

【請求項 1 7】

前記中継装置アドレス取得手段又は補完アドレス取得手段により取得された中継装置のアドレスが正当なアドレスであるかを確認するアドレス確認手段を備え、

10

20

30

40

50

そのアドレス確認手段により前記アドレスが不正なアドレスであると確認された場合に、前記迷惑メール判定手段は、前記送信元から送信された電子メールを迷惑メールであると判定するものであることを特徴とする請求項 1 4 から 1 6 のいずれかに記載の電子メールフィルタリングシステム。

【請求項 1 8】

前記迷惑メール中継確率取得手段は、前記中継装置アドレス取得手段又は補完アドレス取得手段により取得されたアドレスの中継装置に対応する前記メール情報が前記情報記憶手段に不在の新出の中継装置である場合には、その中継装置に対し、前記迷惑メールである確率として所定値を付与することを特徴とする請求項 1 3 から 1 7 のいずれかに記載の電子メールフィルタリングシステム。

10

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、不特定多数の送信元から送信された多数の電子メールの中から迷惑メールをフィルタリングすることが可能な電子メールフィルタリングプログラム、電子メールフィルタリング方法、電子メールフィルタリングシステムに関する。

【背景技術】

【0002】

近年において、ネットワークの発展により、誰しものが気軽に簡単に電子メール（以下、必要に応じて単に「メール」と称する）を送受信できるようになったことに伴い、所謂スパムメール(spam mail)の数も増大している。ここで、「スパムメール」とは、受信者の意図を無視して事前の要請や同意なしに、無差別かつ大量発信されるメールを意味するものである。なお、このスパムメールの同義語として、「迷惑メール」、「ジャンクメール」、「UCE (Unsolicited Commercial Email)」、「UBE (Unsolicited Bulk Email)」などがある。

20

【0003】

このようなスパムメールは、添付ファイルなどによるウイルス感染や、不要なメールの増加による受信者の業務生産性及び効率の低下や、トラフィックの増加によるサーバ及びネットワークへの負荷増大や、詐欺サイトへの誘導などによるプライバシーや機密情報の漏洩などの点において、個人及び団体を問わずに脅威となり得るものである。

30

【0004】

上記のようなスパムメールによる問題は既に社会問題の域にまで達している。メールアドレスが安価に入手可能であることや、定額料金的高速通信が安価で提供されていることなどを鑑みると、今後、スパムメールは減少することなくますます増加していくと考えられ、スパムメールに対する有効な対策が早急に要求されている。

【0005】

現在、使用又は提案されているスパムメール対策としては、送信者を特定する技術や、受信メールをフィルタにかけて選別する方法などがある。

【0006】

40

送信者を特定する技術をスパムメール対策として用いた場合、送信元を特定することによって、差出人を偽るスプーフィングやフィッシング (Phishing) などのメールを受信前に見分けることが可能となる。しかし、その反面で、ドメインを偽装しないスパム業者のメールは排除することができないという問題がある。

【0007】

また、受信メールの選別に用いられるフィルタとして代表的なフィルタとしては、アドレスフィルタやテキストフィルタなどが挙げられる。ここで、アドレスフィルタは、メールに記載されているメールアドレス（例えば、example@xxx.ac.jp）やIPアドレス（例えば、133.aa.bbb.cc）に基づいてメールを選別するフィルタである。

【0008】

50

アドレスフィルタとしては、ブラックリスト(受信拒否リスト)に掲載された送信者からのメールを排除するブラックリストフィルタや、ホワイトリスト(受信許可リスト)に予め登録されることによって明示的に承認された送信者からのメールだけを受信するホワイトリストフィルタなどがある。

【0009】

例えば、特開2001-156834号公報(特許文献1)には、電子メールが到着した際に、その電子メールの差出人を、ユーザによって登録されたホワイトリスト又はブラックリストと対比させ、許可された電子メールのみをファクシミリ装置へ送出可能とするFAXサーバシステムが開示されている。

【0010】

一方で、テキストフィルタは、指定されたヘッダフィールドもしくはヘッダ全体、又は本文に含まれる文字列や文法規則に基づいてメールを選別するフィルタである。このテキストフィルタによれば、例えば、ヘッダフィールドの「Subject:」(題名)に「未承諾広告」などの特定の文字列を検出した場合にスパムメールとして検出するように設定することができる。また、RFC(Request for Comments)の規定に則したヘッダであるかなどの文法規則をチェックした場合に、不適切なものが検出された場合にスパムメールとして検出するように設定することができる。

【0011】

また、最近では、2002年にPaul Grahamによって提案された、ベイズ理論を用いるベイジアンフィルタが有名である。ベイジアンフィルタとは、ベイズ単語分布フィルタとも呼ばれ、スパムメールに出現する単語とハムメールに出現する単語の出現確率の違いを利用したフィルタリング手法である。ベイジアンフィルタは、過去の情報を利用する学習型のフィルタであるので、学習するほど判定精度が向上するフィルタである。

【特許文献1】特開2001-156834号公報

【発明の開示】

【発明が解決しようとする課題】

【0012】

しかしながら、アドレスフィルタは管理(メンテナンス)の作業が煩雑である上に、管理するアドレスの件数の多少に応じて、メールが過剰に受信拒否されたり、逆にスパムメールを容易に通過させたりする事例が生じ得るといった問題点があった。

【0013】

例えば、ブラックリストフィルタは、ブラックリストへの登録や削除などのメンテナンス作業が煩雑である。また、ブラックリストフィルタは、ブラックリストに登録されたメールアドレスのみが受信拒否されるので、メールアドレスが偽装されると、容易にブラックリストフィルタを通過してしまうことになる。

【0014】

また、このブラックリストフィルタが、特定のIPアドレスの範囲にある送信元からの送信であった場合や特定の国の送信元からの送信であった場合に受信を拒否するものであれば、受信すべきハムメールまでも過剰に拒否されかねない。

【0015】

一方で、ホワイトリストフィルタもまた、ホワイトリストへの登録や削除などのメンテナンス作業が煩雑である。また、スパムメールのアドレスがホワイトリストに登録されたメールアドレスと一致するように偽装された場合には、スパムメールがホワイトリストフィルタを通過し、ユーザの元に届いてしまうことになる。さらに、ホワイトリストフィルタが、送信元が特定のIPアドレスの範囲にあるメールの受信を許可するものである場合には、スパムメールを完全に受信拒否することができない。

【0016】

また、テキストフィルタは、フィルタリングのルールを一つ一つ追加(学習)させる作業が煩雑である上に、例えば、ルールとして「未承諾広告」という文字列を追加したとしても、その文字列が「未承諾_広告」のように偽装された場合にはフィルタリングされな

10

20

30

40

50

いなど、判定精度が低いという問題点があった。

【0017】

上述したベイジアンフィルタにもいくつかの問題点がある。例えば、スパム単語データベースの更新に時間がかかることや、学習し続けることでそのデータベースのディスク消費量が他のフィルタよりも大きい。また、販売や広告などのメールを望んで受け取っている場合(オプトイン)であっても、そのメールがスパムメールと認識される確率が非常に高い。さらに、誤認識され易い公告付きメールなどをハムメールとして学習させると、今度はスパムメールがハムメールと認識されてしまうなど、学習サンプルによって判定精度に差が生じる。また、構文解析の難しい日本語で書かれたメールでの判定精度は、英語で書かれたメールに比べて低く、誤判定される傾向にある。

10

【0018】

ここで、ベイジアンフィルタの最も重大な問題点は、正当な送信者に対しても、メールで使用可能な語句を制限してしまうことである。つまり、ベイジアンフィルタは、正当であり重要なメールであっても、スパムメールに度々使用される語句が使用されているメールをスパムメールと認識してしまう。一方で、ハムメールに頻繁に使用される語句を用いてスパムメールを構成した場合には、ベイジアンフィルタはそのメールをスパムメールとして認識しない。

【0019】

上記のようなベイジアンフィルタの問題点を解消する方法としては、アドレスフィルタであるホワイトリストに受信を許可するユーザを登録して、両フィルタを併用する方法があるが、上記したようなホワイトリストの問題点によって十分な効果を上げることができない。

20

【0020】

本発明は、上述した問題点を解決するためになされたものであり、電子メール中に含まれる情報(アドレスやテキストなど)に依存することなく、配信経路上における中継装置が過去に中継した迷惑メール(スパムメール)及び正当なメール(ハムメール)の各頻度に基づいて、不特定多数の送信元から送信された多数の電子メールの中から迷惑メールをフィルタリングすることが可能な電子メールフィルタリングプログラム、電子メールフィルタリング方法、電子メールフィルタリングシステムを提供することを目的としている。

【課題を解決するための手段】

30

【0021】

この目的を達成するために、請求項1記載の電子メールフィルタリングプログラムは、送信元から送信された不特定多数の電子メールに含まれる迷惑メールのフィルタリングを制御装置に実行させるためのプログラムであり送信元から送信された電子メールのヘッダ情報を参照して、その電子メールの配信経路上における少なくとも1の中継装置のアドレスを取得する中継装置アドレス取得ステップと、前記配信経路上における中継ルータのアドレスを取得する中継ルータアドレス取得ステップと、その中継ルータアドレス取得ステップにより得られた中継ルータのアドレスに基づいて、前記中継装置アドレス取得ステップにおいてアドレスが取得された中継装置以外の中継装置であって前記配信経路を補完する中継装置のアドレスを取得する補完アドレス取得ステップと、前記中継装置アドレス取得ステップにおいてアドレスが取得された中継装置と、前記補完アドレス取得ステップにおいてアドレスが取得された中継装置とに対し、それらの中継装置が過去に中継した迷惑メール及び正当なメールの各頻度を示す情報であって情報記憶手段に記憶されているメール情報に基づいて、該中継装置によって中継された電子メールが迷惑メールである確率を、ベイズ確率モデルを用いて得る迷惑メール中継確率取得ステップと、その迷惑メール中継確率取得ステップにおいて得られた確率に基づいて、前記送信元から送信された電子メールが迷惑メールである確率を得る迷惑メール受信確率取得ステップと、その迷惑メール受信確率取得ステップにおいて得られた確率に応じて前記電子メールを所定区分に分類するメール判定ステップとを備えている。

40

50

【 0 0 2 2 】

請求項1記載の電子メールフィルタリングプログラムによれば、まず、中継装置アドレス取得ステップにより、送信元から送信された電子メールのヘッダ情報を参照して、その電子メールの配信経路上における少なくとも1の中継装置のアドレスが取得される。一方で、中継ルータアドレス取得ステップにより、配信経路上における中継ルータのアドレスが得られ、その中継ルータのアドレスに基づいて、補完アドレス取得ステップにより、中継装置アドレス取得ステップによってアドレスが取得された中継装置以外の中継装置であって、配信経路を補完する中継装置のアドレスが得られる。次に、迷惑メール中継確率取得ステップにより、中継装置アドレス取得ステップ又は補完アドレス取得ステップによってアドレスの取得された中継装置に対し、情報記憶手段に記憶されているメール情報に基づいて、該中継装置によって中継された電子メールが迷惑メールである確率が得られる。ここで、この迷惑メール中継確率取得ステップにより得られる確率は、ベイズ確率モデルを用いて得られる確率である。

10

【 0 0 2 3 】

次に、迷惑メール受信確率取得ステップにより、迷惑メール中継確率取得ステップによって得られた確率に基づいて、送信元から送信された電子メールが迷惑メールである確率が得られる。そして、迷惑メール受信確率取得ステップによって得られた送信元から送信された電子メールが迷惑メールである確率に応じて、その電子メールが、メール判定ステップによって判定されて所定区分に分類される。

20

【 0 0 2 4 】

即ち、請求項1記載の電子メールフィルタリングプログラムによれば、電子メールの配信経路上の中継装置、及び、配信経路を補完する中継装置をアドレスによって特定した上で、メール情報に基づき、送信元から送信された電子メールを、例えば、迷惑メールであるとして分類することができる。

【 0 0 2 5 】

なお、特許請求の範囲における用語「迷惑メール」とは、受信者の意図を無視して事前の要請や同意なしに、無差別かつ大量発信されるメール(所謂「スパムメール」)を意味する。また、特許請求の範囲における用語「正当なメール」とは、「迷惑メール」の対義語であり、迷惑メールでないメール、即ち、送信者と受信者との間で互いにそのメールの受け渡しを行う必然性のあるメール(所謂「ハムメール」)を意味する。また、特許請求の範囲における用語「メール情報」とは、中継装置が過去に中継した迷惑メール及び正当なメールの各頻度を示す情報を意味する。また、特許請求の範囲における用語「中継装置」とは、配信経路上においてメールが経由する装置及びメールが経由する可能性のある(経由すると推定される)装置を意味する。よって、特許請求の範囲における「中継装置」には、メールを中継する中継サーバだけでなく、メールの送信元となる端末や、配信経路追跡を行った場合にメールを中継したと推定される装置も含まれる。

30

【 0 0 2 6 】

請求項2記載の電子メールフィルタリングプログラムは、請求項1記載の電子メールフィルタリングプログラムにおいて、前記メール判定ステップは、前記迷惑メール受信確率取得ステップにおいて得られた確率が、第1閾値を越えたか又は第1閾値以上であった場合に、前記送信元から送信された電子メールを迷惑メールであると判定する迷惑メール判定ステップと、前記迷惑メール受信確率取得ステップにおいて得られた確率が、第2閾値未満又は第2閾値以下の場合に、前記送信元から送信された電子メールを正当なメールであると判定する正当メール判定ステップとを含み、前記迷惑メール判定ステップ又は前記正当メール判定ステップにおいて判定された結果に応じて、前記中継装置アドレス取得ステップ又は補完アドレス取得ステップにおいて取得されたアドレスの中継装置について、前記情報記憶手段に記憶されている前記メール情報を更新する情報更新ステップを備えている。

40

50

【 0 0 2 7 】

請求項 2 記載の電子メールフィルタリングプログラムによれば、請求項 1 記載の電子メールフィルタリングプログラムと同様に作用する上、メール判定ステップは、迷惑メール判定ステップと正当メール判定ステップとを含んでいる。即ち、迷惑メール受信確率取得ステップによって得られた確率が、第 1 閾値を越えたか又は第 1 閾値以上であった場合には、迷惑メール判定ステップにより、送信元から送信された電子メールが迷惑メールであると判定される。一方で、迷惑メール受信確率取得ステップによって得られた確率が、第 2 閾値未満又は第 2 閾値以下の場合には、正当メール判定ステップにより、送信元から送信された電子メールが正当なメールであると判定される。

10

【 0 0 2 8 】

そして次に、情報更新ステップによって、上記の迷惑メール判定ステップ又は正当メール判定ステップの実行によって判定された結果に応じて、情報記憶手段に記憶されているメール情報が更新される。即ち、中継装置アドレス取得ステップによってアドレスが取得された中継装置に対応する情報記憶手段のメール情報が、上記の迷惑メール判定ステップ又は正当メール判定ステップによって判定された結果に応じて更新される。

【 0 0 2 9 】

請求項 3 記載の電子メールフィルタリングプログラムは、請求項 2 記載の電子メールフィルタリングプログラムにおいて、前記メール判定ステップは、前記第 1 閾値と前記第 2 閾値とが異なる場合に、前記迷惑メール判定ステップ又は前記正当メール判定ステップにおいて前記迷惑メール又は前記正当なメールのいずれにも非該当であると判定された前記電子メールを、不確定メールと認識する不確定メール認識ステップをさらに含む。

20

【 0 0 3 0 】

請求項 3 記載の電子メールフィルタリングプログラムによれば、請求項 2 記載の電子メールフィルタリングプログラムと同様に作用する上、メール判定ステップは、不確定メール認識ステップを含んでいる。即ち、第 1 閾値と第 2 閾値とが異なる場合に、迷惑メール判定ステップ又は正当メール判定ステップによって電子メールが迷惑メールでも正当なメールでもないとして判定された場合には、不確定メール認識ステップにより、迷惑メールでも正当なメールでもないとして判定された電子メールが不確定メールとして認識される。

【 0 0 3 1 】

請求項 4 記載の電子メールフィルタリングプログラムは、請求項 3 記載の電子メールフィルタリングプログラムにおいて、前記メール判定ステップは、前記不確定メール認識ステップにおいて不確定メールと認識された前記電子メールに対し、その電子メールに含まれるテキスト情報を利用することによって、その電子メールが迷惑メールであるか又は正当なメールであるかを判定する不確定メール再判定ステップをさらに含む。

30

【 0 0 3 2 】

請求項 4 記載の電子メールフィルタリングプログラムによれば、請求項 3 記載の電子メールフィルタリングプログラムと同様に作用する上、メール判定ステップは、不確定メール再判定ステップを含んでいる。即ち、不確定メールに認識ステップによって電子メールが不確定メールと認識された場合には、不確定メール再判定ステップにより、その不確定メールに含まれるテキスト情報に基づいて、その不確定メールが迷惑メールであるか又は正当なメールであるかが判定される。

40

【 0 0 3 3 】

なお、特許請求の範囲における「不確定メールに含まれるテキスト情報」との記載は、本文テキストやメールヘッダにおけるメールアドレスなど、不確定メールの中でテキストによって記述された情報を全て包含している。

【 0 0 3 4 】

請求項 5 記載の電子メールフィルタリングプログラムは、請求項 2 から 4 のいずれかに記載の電子メールフィルタリングプログラムにおいて、前記中継装置アドレス取得ステップ又は補完アドレス取得ステップにより取得された中継装置のアドレスが正当なアドレス

50

であるかを確認するアドレス確認ステップを備えていると共に、そのアドレス確認ステップにより前記アドレスが不正なアドレスであると確認された場合には、前記メール判定ステップにおいて、前記送信元から送信された電子メールが迷惑メールであると判定する。

【 0 0 3 5 】

請求項 5 記載の電子メールフィルタリングプログラムによれば、請求項 2 から 4 のいずれかに記載の電子メールフィルタリングプログラムと同様に作用する上、アドレス確認ステップにより、中継装置アドレス取得ステップ又は補完アドレス取得ステップによって取得された中継装置のアドレスが正当なアドレスであるかが確認される。その結果、アドレスが不正なアドレスであると確認された場合には、メール判定ステップにより、送信元から送信された電子メールが迷惑メールであると判定される。

10

【 0 0 3 6 】

請求項 6 記載の電子メールフィルタリングプログラムは、請求項 1 から 5 のいずれかに記載の電子メールフィルタリングプログラムにおいて、前記中継装置アドレス取得ステップ又は補完アドレス取得ステップにより取得されたアドレスの中継装置に対応する前記メール情報が前記情報記憶手段に不在の新出の中継装置である場合には、前記迷惑メール中継確率取得ステップにおいて、その新出の中継装置に対し、前記迷惑メールである確率として所定値を付与する。

20

【 0 0 3 7 】

請求項 6 記載の電子メールフィルタリングプログラムによれば、請求項 1 から 5 のいずれかに記載の電子メールフィルタリングプログラムと同様に作用する上、中継装置アドレス取得ステップ又は補完アドレス取得ステップによって取得されたアドレスの中継装置に対応するメール情報が情報記憶手段に記憶されていない新出の中継装置である場合には、迷惑メール中継確率取得ステップにより、その新出の中継装置に対し、迷惑メールである確率として所定値が付与される。

【 0 0 4 1 】

請求項 7 記載の電子メールフィルタリング方法は、送信元から送信された不特定多数の電子メールに含まれる迷惑メールをフィルタリングすることができる方法であって、送信元から送信された電子メールのヘッダ情報を参照して、その電子メールの配信経路上における少なくとも 1 の中継装置のアドレスを取得する中継装置アドレス取得手段と、前記配信経路上における中継ルータのアドレスを取得する中継ルータアドレス取得手段と、その中継ルータアドレス取得手段により得られた中継ルータのアドレスに基づいて、前記中継装置アドレス取得手段によりアドレスが取得された中継装置以外の中継装置であって前記配信経路を補完する中継装置のアドレスを取得する補完アドレス取得手段と、前記中継装置アドレス取得手段によりアドレスが取得された中継装置と、前記補完アドレス取得手段によりアドレスが取得された中継装置とに対し、それらの中継装置が過去に中継した迷惑メール及び正当なメールの各頻度を示す情報であって情報記憶手段に記憶されているメール情報に基づいて、該中継装置によって中継された電子メールが迷惑メールである確率を、ベイズ確率モデルを用いて得る迷惑メール中継確率取得手段と、その迷惑メール中継確率取得手段により得られた確率に基づいて、前記送信元から送信された電子メールが迷惑メールである確率を得る迷惑メール受信確率取得手段と、その迷惑メール受信確率取得手段により得られた確率に応じて前記電子メールを所定区分に分類するメール判定手段とを備えていることを特徴とする電子メールフィルタリング方法。

30

40

【 0 0 4 2 】

請求項 8 記載の電子メールフィルタリング方法は、請求項 7 記載の電子メールフィルタリング方法において、前記メール判定手段は、前記迷惑メール受信確率取得手段により得

50

られた確率が、第1閾値を越えたか又は第1閾値以上であった場合に、前記送信元から送信された電子メールを迷惑メールであると判定する迷惑メール判定手段と、前記迷惑メール受信確率取得手段により得られた確率が、第2閾値未満又は第2閾値以下の場合に、前記送信元から送信された電子メールを正当なメールであると判定する正当メール判定手段とを備えており、前記迷惑メール判定手段又は前記正当メール判定手段により判定された結果に応じて、前記中継装置アドレス取得手段又は補完アドレス取得手段により取得されたアドレスの中継装置について、前記情報記憶手段に記憶されている前記メール情報を更新する情報更新手段を備えている。

【0043】

請求項9記載の電子メールフィルタリング方法は、請求項8記載の電子メールフィルタリング方法において、前記メール判定手段は、前記第1閾値と前記第2閾値とが異なる場合に、前記迷惑メール判定手段又は前記正当メール判定手段により前記迷惑メール又は前記正当なメールのいずれにも非該当であると判定された前記電子メールを、不確定メールと認識する不確定メール認識手段をさらに備えている。

10

【0044】

請求項10記載の電子メールフィルタリング方法は、請求項9記載の電子メールフィルタリング方法において、前記メール判定手段は、前記不確定メール認識手段により不確定メールと認識された前記電子メールに対し、その電子メールに含まれるテキスト情報を利用することによって、その電子メールが迷惑メールであるか又は正当なメールであるかを判定する不確定メール再判定手段をさらに備えている。

20

【0045】

請求項11記載の電子メールフィルタリング方法は、請求項8から10のいずれかに記載の電子メールフィルタリング方法において、前記中継装置アドレス取得手段又は補完アドレス取得手段により取得された中継装置のアドレスが正当なアドレスであるかを確認するアドレス確認手段を備え、そのアドレス確認手段により前記アドレスが不正なアドレスであると確認された場合に、前記迷惑メール判定手段は、前記送信元から送信された電子メールを迷惑メールであると判定する。

30

【0046】

請求項12記載の電子メールフィルタリング方法は、請求項7から11のいずれかに記載の電子メールフィルタリング方法において、前記迷惑メール中継確率取得手段は、前記中継装置アドレス取得手段又は補完アドレス取得手段により取得されたアドレスの中継装置に対応する前記メール情報が前記情報記憶手段に不在の新出の中継装置である場合には、その中継装置に対し、前記迷惑メールである確率として所定値を付与する。

【0048】

請求項13記載の電子メールフィルタリングシステムは、電子メールを伝送可能な経路上において、送信元から送信された不特定多数の電子メールに含まれる迷惑メールをフィルタリングすることが可能なシステムであって、送信元から送信された電子メールのヘッダ情報を参照して、その電子メールの配信経路上における少なくとも1の中継装置のアドレスを取得する中継装置アドレス取得手段と、1の中継装置に対し、その中継装置によって過去に中継された迷惑メール及び正当なメールの各頻度を示すメール情報を記憶する情報記憶手段と、前記配信経路上における中継ルータのアドレスを取得する中継ルータアドレス取得手段と、その中継ルータアドレス取得手段により得られた中継ルータのアドレスに基づいて、前記中継装置アドレス取得手段においてアドレスが取得された中継装置以外の中継装置であって前記配信経路を補完する中継装置のアドレスを取得する補完アドレス取得手段と、前記中継装置アドレス取得手段によりアドレスが取得された中継装置と、前記

40

50

補完アドレス取得手段によりアドレスが取得された中継装置とに対し、情報記憶手段に記憶されている前記メール情報に基づいて、それらの中継装置によって中継された電子メールが迷惑メールである確率を、ベイズ確率モデルを用いて得る迷惑メール中継確率取得手段と、その迷惑メール中継確率演算手段により得られた確率に基づいて、前記送信元から送信された電子メールが迷惑メールである確率を得る迷惑メール受信確率取得手段と、その迷惑メール受信確率取得手段により得られた確率に応じて前記電子メールを所定区分に分類するメール判定手段とを備えている。

【 0 0 4 9 】

請求項 1 4 記載の電子メールフィルタリングシステムは、請求項 1 3 記載の電子メールフィルタリングシステムにおいて、前記メール判定手段は、前記迷惑メール受信確率取得手段により得られた確率が、第 1 閾値を越えたか又は第 1 閾値以上であった場合に、前記送信元から送信された電子メールを迷惑メールであると判定する迷惑メール判定手段と、前記迷惑メール受信確率取得手段により得られた確率が、第 2 閾値未満又は第 2 閾値以下の場合に、前記送信元から送信された電子メールを正当なメールであると判定する正当メール判定手段とを備えており、前記迷惑メール判定手段又は前記正当メール判定手段による判定結果に応じて、前記中継装置アドレス取得手段又は補完アドレス取得手段により取得されたアドレスの中継装置について、前記情報記憶手段に記憶されている前記メール情報を更新する情報更新手段を備えている。

10

20

【 0 0 5 0 】

請求項 1 5 記載の電子メールフィルタリングシステムは、請求項 1 4 記載の電子メールフィルタリングシステムにおいて、前記メール判定手段は、前記第 1 閾値と前記第 2 閾値とが異なる場合に、前記迷惑メール判定手段又は前記通常メール判定手段による判定が前記迷惑メール又は前記正当なメールのいずれにも非該当である電子メールを、不確定メールと認識する不確定メール認識手段をさらに備えている。

【 0 0 5 1 】

請求項 1 6 記載の電子メールフィルタリングシステムは、請求項 1 5 記載の電子メールフィルタリングシステムにおいて、前記メール判定手段は、前記不確定メール認識手段において不確定メールと認識された前記電子メールに対し、その電子メールに含まれるテキスト情報を利用することによって、その電子メールが迷惑メールであるか又は正当なメールであるかを判定する不確定メール再判定手段をさらに備えている。

30

【 0 0 5 2 】

請求項 1 7 記載の電子メールフィルタリングシステムは、請求項 1 4 から 1 6 のいずれかに記載の電子メールフィルタリングシステムにおいて、前記中継装置アドレス取得手段又は補完アドレス取得手段により取得された中継装置のアドレスが正当なアドレスであることを確認するアドレス確認手段を備え、そのアドレス確認手段により前記アドレスが不正なアドレスであると確認された場合に、前記迷惑メール判定手段は、前記送信元から送信された電子メールを迷惑メールであると判定する。

40

【 0 0 5 3 】

請求項 1 8 記載の電子メールフィルタリングシステムは、請求項 1 3 から 1 7 のいずれかに記載の電子メールフィルタリングシステムにおいて、前記迷惑メール中継確率取得手段は、前記中継装置アドレス取得手段又は補完アドレス取得手段により取得されたアドレスの中継装置に対応する前記メール情報が前記情報記憶手段に不在の新出の中継装置である場合には、その中継装置に対し、前記迷惑メールである確率として所定値を付与する。

【 発明の効果 】

50

【 0 0 5 5 】

本発明の電子メールフィルタリングプログラム、電子メールフィルタリング方法、電子メールフィルタリングシステムによれば、電子メールのヘッダ情報における記述に基づいて配信経路上の中継装置をアドレスによって特定した上で、メール情報に基づき、送信元から送信された電子メールを、例えば、1区分として迷惑メールであると区分して分類することができる。

【 0 0 5 6 】

ここで、配信経路上における中継装置が過去に迷惑メールを中継した頻度の多さは、その中継装置が迷惑メールの配信経路上にある可能性の高さに対応する。即ち、ある中継装置に対応するメール情報が、過去における迷惑メールの中継頻度が高いことを示すものであれば、その中継装置は、悪質な送信者から送信された迷惑メールの配信経路上の中継装置である可能性が高い。

10

【 0 0 5 7 】

よって、本発明の電子メールフィルタリングプログラム、電子メールフィルタリング方法、電子メールフィルタリングシステムによれば、配信経路の中継装置をアドレスによって特定した上で、その特定された中継装置が過去に中継した迷惑メール及び正当なメールの頻度を示すメール情報をメール判定のために利用して、受信したメールを所定区分、例えば、迷惑メールとして分類することができる。また、その一方で、正当なメールが過剰に拒否されることを抑制できるという効果がある。このように、迷惑メールが確実に検出されると、その結果として、迷惑メールと判定されたメールを削除する、又は、受信を拒否するなどの処理を施すことが可能となり、ユーザが迷惑メールにより受ける実害を低減できる。

20

【 0 0 5 8 】

また、迷惑メールにおいて、配信経路における中継装置のアドレス（IPアドレス、メールアドレス）が偽装される場合、同じアドレスが複数回使われることが少なく、結果として、偽装されたアドレスは1回のみ出現となる。そのため、単語の出現回数から判定する従来のベイズフィルタでは、そのような1回きりの偽装を迷惑メールとして判定することが困難である。

【 0 0 5 9 】

一方で、本発明では、中継装置が過去に中継したメール情報、即ち、迷惑メールの頻度と正当なメールの頻度とを組合わせてベイズ理論を適用することによって、迷惑メールであるか否かの判定力を向上させることができるという効果がある。

30

【 0 0 6 0 】

また、中継装置が過去に中継した迷惑メール及び正当なメールの頻度を示すメール情報をメール判定のために利用するので、従来のテキストフィルタのような膨大なデータの蓄積を必要とせず、データベースによる記憶装置（メモリやディスクなど）の消費量を抑制することができるという効果がある。また、送信者による偽装が容易である電子メールに含まれるテキストを利用することなくメールを区分できるので、例えば、所定の閾値を境界としてそれ以上又はそれを越えた場合にそのメールを迷惑メールであると区分することにより、迷惑メールを確実に検出できると共に、正当なメールが過剰に拒否されることを抑制できるという効果がある。

40

【 0 0 6 1 】

さらに、送信者による偽装が容易である送信者のアドレス（メールアドレス）を利用することなくメールを区分できるので、例えば、所定の閾値を境界としてそれ以上又はそれを越えた場合にそのメールを迷惑メールであると区分することにより、迷惑メールを確実に検出できると共に、正当なメールが過剰に拒否されることを抑制できるという効果がある。さらに、従来のアドレスフィルタのように、悪質な送信者との齟齬ごっこのような偽装アドレスの登録及び削除を繰り返す必要がなくなり、管理の負担が軽減されるという効果もある。

【 0 0 6 2 】

50

加えて、配信経路上の中継装置のIPアドレスを利用してメールの判定を行うので、その判定結果は言語情報に依存しない。よって、構文解析が難しく従来のテキストフィルタでは誤判定されやすかった日本語のメールであっても、迷惑メールであるか正当なメールであるかを確実に検出することができるという効果がある。

【0063】

また、本発明の電子メールフィルタリングプログラム、電子メールフィルタリング方法、電子メールフィルタリングシステムによれば、迷惑メール又は正当なメールであると判定されると、その結果に基づいて、該メールの配信経路上の中継装置に対応する情報記憶手段に記憶されているメール情報が更新される。即ち、判定結果を学習するので、学習するほどその判定精度が向上するという効果がある。

10

【0064】

さらに、本発明の電子メールフィルタリングプログラム、電子メールフィルタリング方法、電子メールフィルタリングシステムによれば、迷惑メールとも正当なメールとも判定がつかず不確定要素の高いメールが不確定メール（グレイメール）として認識される。このように曖昧な分類区分を設けることにより、正当なメールが迷惑メールとして、又は、迷惑メールが正当なメールとして誤判定されることを防止できるという効果がある。その結果として、そのような誤判定に基づいてユーザが被り得る問題を回避することができるという効果がある。

【0065】

加えて、本発明の電子メールフィルタリングプログラム、電子メールフィルタリング方法、電子メールフィルタリングシステムによれば、不確定メールであると認識されたメールに対し、そのメールに含まれるテキスト情報に基づいて迷惑メールであるか正当なメールであるかの判定が行われるので、配信経路上を通過して配信される電子メールが迷惑メールである場合にそれを確実に検出できると共に、正当なメールが過剰に拒否されることを抑制できるという効果がある。

20

【0066】

また、本発明の電子メールフィルタリングプログラム、電子メールフィルタリング方法、電子メールフィルタリングシステムによれば、電子メールのヘッダ情報に基づいて特定された中継装置が正当に登録されているアドレスであるかを確認し、その際、不正なアドレスであれば、そのメールが迷惑メールであると判定される。よって、ヘッダ情報が明らかに不正であることが確認されれば迷惑メールとして処理されるので、処理を効率化できると共に、迷惑メールであるか又は正当なメールであるかの判定精度を向上させることができるという効果がある。

30

【0067】

また、本発明の電子メールフィルタリングプログラム、電子メールフィルタリング方法、電子メールフィルタリングシステムによれば、情報記憶手段にメール情報の記憶されていない新出の中継装置に対しては、そのメールが迷惑メールである確率として所定値を用いるので、幅広い配信経路に対して適用可能であると共に、新たな配信経路が生じやすい初見メールに対しても対応可能であるという効果がある。

【0068】

さらに、本発明の電子メールフィルタリングプログラム、電子メールフィルタリング方法、電子メールフィルタリングシステムによれば、中継ルータのアドレスを取得し、その中継ルータのアドレスに基づいて、ヘッダ情報を参照することによってアドレスが取得されていた中継装置以外の中継装置が、電子メールの配信経路上に補完される。よって、迷惑メールであるかの判定のために利用する中継装置の数が増加するので、同一の経路を経由する重なりを増やすことができる。その結果として、迷惑メールが配信される傾向にある配信経路と正当なメールが配信される傾向にある配信経路との区別をより明確にすることができるので、迷惑メールであるか又は正当なメールであるかの判定精度を向上させることができるという効果がある。

40

【発明を実施するための最良の形態】

50

【0069】

以下、本発明の好ましい実施例について、添付図面を参照しつつ説明する。まず、本発明の電子メールフィルタリングシステムの第1実施例について説明する。図1は、本発明の第1実施例における電子メールフィルタリングシステムが実装される電子メールの配信経路の一例を示す模式図である。図1に示すように、6つの端末Ts (Ts₁ ~ Ts₆) を送信側の端末とした場合には、それらの端末から受信側の端末Trへ向けて送信されたメールは、それらの間に介在された5個のメールサーバS (S₁ ~ S₅)により中継されて、端末Trへ配信される。この場合、メールの配信経路は、図1に示す通り、送信側端末Tsの数の、即ち、6通り存在することになる。

【0070】

具体的には後で詳述するが、本実施例の電子メールフィルタリングシステムでは、配信経路を経由したメールの履歴に基づいて、配信経路における「スパムメールが中継される傾向」であるか「ハムメールが中継される傾向」であるかを推測し、そのような推測に基づいて、送信側端末Ts (図1では端末Ts₁ ~ Ts₆)から送信されたメールがスパムメールであるか否かを判定する。

【0071】

つまり、図1に示す例でいえば、6つの配信経路上においてメールが受信側端末Trに到達するまでに経由する装置(以下、「中継装置」と称する)、即ち、送信側端末Ts (Ts₁ ~ Ts₆)及びメールサーバS (S₁ ~ S₅)のそれぞれについて、過去にスパムメールが通過した頻度の情報及びハムメールが通過した頻度の情報を記憶しておく。次いで、その情報に基づいて、ある配信経路を通るメールがスパムメールである確率を、ベイズ確率モデルを用いて算出する。そして、そのように算出された確率に基づいて、その配信経路を通るメールがスパムメールであるか否かを判定する。

【0072】

例えば、送信側端末Ts₁が受信側端末Trに対しスパムメールのみを送信していた場合には、メールサーバS₁, S₄, S₅を含む経路Aを共有する送信側端末Ts₂が受信側Trへ最初のメール(以下、ある端末へ初めて送信したメールを「初見メール」と称する)を送信した場合に、そのメールを「スパムメールらしい」と推測することができる。同様に、例えば、送信側端末Ts₅が受信側端末Trに対しハムメールのみを送信していた場合には、メールサーバS₃, S₄, S₅を含む経路Cを共有する送信側端末Ts₆から受信側端末Trへ向けて送信されたメールは、初見メールであっても「ハムメールらしい」と推測することができる。

【0073】

送信側端末Ts₁ ~ Ts₆から受信側端末Trへメールが配信される際に経由するメールサーバSは、そのメールにおけるヘッダ情報を参照することにより特定することができる。ヘッダ情報の1つである「Received:」フィールドは、メールが、送信側の端末Ts (図1では端末Ts₁ ~ Ts₆)から送信されてから、受信側の端末Trに到達するまでに経由する(中継される)メールサーバS (図1ではメールサーバS₁ ~ S₅)を示すフィールド(情報)である。

【0074】

ここで、図2を参照して、「Received:」フィールドに記録されている情報について説明する。図2は、メールに付加された「Received:」フィールドを示す模式図である。ここで、メールサーバSは、上記のように「Received:」フィールドへ情報を追加することは許可されているが、既に存在する「Received:」フィールドの情報を消去や変更することは禁止されている。

【0075】

よって、送信側端末Tsから受信側端末Trへ到達するまでに経由する(中継される)メールサーバSの数の多いほど、「Received:」フィールドの数は増えることになる。図2に示す例では、情報R1及び情報R2の2つの「Received:」フィールドが存在している。即ち、図2に示す「Received:」フィールドが記録されたメールは、2つのメールサ

10

20

30

40

50

サーバによって中継されたことを示す。

【0076】

「Received:」フィールドは、図2に示すように、このフィールドの開始文字列である「Received:」が必須で記録され、この文字列「Received:」以下に、文字列「from」で始まる送信ホストの情報や、文字列「by」で始まる受信ホストの情報などが任意に記憶される。ここで、文字列「from」以下に記録される送信ホストの情報は、多くの場合、「送信ホスト名（FQDN名 [IPアドレス]）」の書式で記載されている。なお、「FQDN」とは、「Fully Qualified Domain Name」の略である。一方、文字列「by」以下に記録される受信ホストの情報は、多くの場合、「FQDN名（付加情報）」の書式で記載されている。

10

【0077】

また、「Received:」フィールドが追加される場合には、下から上に向かって追加されるので、下に行くほど送信側端末Tsに近い情報となる。よって、この「Received:」フィールドを遡ることによって、メールの配信経路を受信者側端末Trから送信者側端末Tsまで遡ることができる。

【0078】

よって、図2に示す「Received:」フィールドの記録は、送信者側端末から送信されたメールが、『(1)送信者側端末(2)FQDN名が「ceres.xxx.ne.jp」であり、IPアドレスが「211.6.xxx.78」であるメールサーバ(3)FQDN名が「mx1.xxx.or.jp」であるメールサーバ(4)受信側端末』の配信経路を通して、受信側端末に到達していることを示している。

20

【0079】

ここで、上記したように、「Received:」フィールドはメールを受信したメールサーバSによって追加される。よって、送信者により送信側端末Tsのメールアドレスやヘッダ情報が偽装されていたとしても、その偽装されたメールを受信したメールサーバSは、偽装されたメールアドレスやヘッダ情報とは無関係に、そのメールをどのホスト(IPアドレス)からいつ受信したかを「Received:」フィールドに記録する。このように、「Received:」フィールドに記録される情報は、そのメールサーバSが正規に管理されているものである場合には、送信者による偽装が困難な情報である。

【0080】

本実施例の電子メールフィルタリングシステムは、受信側端末Trに電子メールフィルタリングプログラムを実装することにより、受信側端末Trにおいて、受信したメールがスパムメールであるか否かを判定するものである。

30

【0081】

ここで、図3を参照して、本実施例の電子メールフィルタリングシステムを機能させる受信側端末Trの構成について説明する。図3は、受信側端末Trの構成を示すブロック図である。図3に示すように、受信側端末Trは、受信側端末Tr全体の動作を制御するCPU10と、そのCPU10により実行される制御プログラム12aや固定値データを記憶するROM12と、CPU10により実行される各種処理に必要なデータやプログラム等を一時的に記憶するためのメモリであるRAM14と、記憶部16と、公衆通信網などの通信回線を介してメールサーバSと接続するためのインターフェイス18(I/F18)とを主に備えており、これらの構成がバスライン20によって互いに接続されている。

40

【0082】

ここで、ROM12に格納されている制御プログラム12aには、本実施例の電子メールフィルタリングシステムを機能させる電子メールフィルタリングプログラムが含まれている。この制御プログラム12aに含まれる電子メールフィルタリングプログラムによって実行される処理については後述する。

【0083】

RAM14は、受信メールメモリ14aと、経路情報メモリ14bと、配信木情報メモ

50

リ14cとを備えている。なお、これらのメモリ14a~14cは、いずれも受信側端末Trの電源投入時に初期化される。

【0084】

受信メールメモリ14aは、受信したメールを一時的に記憶するメモリであり、この受信メールメモリ14aに記憶されたメールに対し、後述する電子メールフィルタリングプログラムに従う処理が実行される。

【0085】

経路情報メモリ14bは、送信側端末Tsから送信されたメールが受信側装置Trに受信されるまでの配信経路上の中継装置（送信側端末Ts及びメールサーバS）のIPアドレスを記憶するメモリである。なお、この経路情報メモリ14bに記憶されるメールサーバSのIPアドレスは、受信したメール（受信メールメモリ14aに記憶されたメール）の「Received:」に記録されているIPアドレスだけでなく、後述する中継ルータの探索に基づいて取得されたメールサーバのIPアドレスも含まれる。

10

【0086】

配信木情報メモリ14cは、経路情報メモリ14bに記憶されたIPアドレスに対応する中継装置（送信側端末Ts及びメールサーバS）について、その中継装置が過去に中継したスパムメール及びハムメールの数を後述する中継装置メモリ16aから読み出して一時的に記憶するためのメモリである。

【0087】

記憶部16は、書き換え可能な大容量の記憶装置であり、電源断後もデータを保持する不揮発性のメモリであるハードディスクなどの書き換え可能な不揮発性メモリである。この記憶部16は、中継装置メモリ16aと、スパム受信カウンタ16bと、ハム受信カウンタ16cとを備えている。

20

【0088】

中継装置メモリ16aは、中継装置、即ち、送信側端末Ts（図1ではTs₁~Ts₆）及びメールサーバS（図1ではS₁~S₅）が過去に中継したスパムメール及びハムメールの数を記憶するメモリである。この中継装置メモリ16aは、第1中継装置メモリ16a₁から第n中継装置メモリ16a_nまでのn個のメモリから構成されており、その数（n個）は、過去に受信側端末Trが受信したメールの配信経路上の中継装置（送信側端末Ts及びメールサーバS）として検出された全ての中継装置の数に対応する。

30

【0089】

第1~第n中継装置メモリ16a₁~16a_nは、それぞれ、スパム中継カウンタ16a₁₁~16a_{n1}とハム中継カウンタ16a₁₂~16a_{n2}とを備えている。ここで、スパム中継カウンタ16a₁₁~16a_{n1}は、それぞれ、対応する中継装置がスパムメールを中継（経由）した数を計数するカウンタであり、一方で、ハム中継カウンタ16a₁₂~16a_{n2}は、対応する中継装置がハムメールを中継した数を計数するカウンタである。

【0090】

なお、これらのスパム中継カウンタ16a₁₁~16a_{n1}及びハム中継カウンタ16a₁₂~16a_{n2}の値は、後述する学習処理（図9参照）において、新規の中継装置メモリ16aが作成された場合に、そのメールに対するメール判定処理（図7参照）による判定結果に基づいて、初期値として「0」又は「1」が設定される。

40

【0091】

スパム受信カウンタ16bは、受信側端末Trが受信したスパムメールの数を計数するカウンタであり、ハム受信カウンタ16cは、受信側端末Trが受信したハムメールの数を計数するカウンタである。

【0092】

なお、1の中継装置メモリ16a_x（x=1~n）に対応するスパム中継カウンタ16a_{x1}及びハム中継カウンタ16a_{x2}（x=1~n）の値と、スパム受信カウンタ16b及びハム受信カウンタ16cの値とから得られる値が、その中継装置メモリ16a_xに

50

対応する中継装置（メールサーバS及び送信側端末Ts）の「メール情報」である。

【0093】

次に、上記のように構成された受信側端末Trに実装された電子メールフィルタリングプログラムによって実行される各処理について説明する。図4は、電子メールフィルタリングプログラムによって実行されるメール受信処理を示すフローチャートである。図4に示すメール受信処理は、受信側端末Trにおいて、ユーザがメール受信の指示を行った場合に起動する処理である。

【0094】

図4に示すように、メール受信処理は、まず、その受信側端末Trに接続されるメールサーバS（図1に示すサーバS₅）にメールがあるかを確認し（S1）、メールがあれば（S1:Yes）、そのメールサーバS₅から1のメールを受信し、受信メールメモリ14aにその受信メールのデータを記憶する（S2）。

10

【0095】

S2の処理後、その受信メールにおけるメールヘッダ（「Received:」フィールド）を参照することによって、メールの配信経路の経路情報を得る経路情報取得処理を実行し（S3）、その経路情報取得処理（S3）の処理の結果として取得された経路情報に基づいて、1の配信経路を表す配信木と呼ばれる木構造を構築する配信木構築処理を実行する（S4）。

【0096】

配信木構築処理（S4）の実行後、構築された配信木によって表される1の配信経路を通るメールがスパムメールであるか否かを判定するメール判定処理（S5）を実行する。そして、メール判定処理（S5）の実行後、その判定結果に基づいて、記憶部16に記憶される「メール情報」の更新を行う学習処理（S6）を実行する。なお、これらの経路情報取得処理（S3）、配信木構築処理（S4）、メール判定処理（S5）、学習処理（S6）における具体的な処理については、それぞれ、図5～図7、図9のフローチャートを参照しつつ後述する。

20

【0097】

そして、学習処理（S6）の実行後、メールサーバS₅に受信すべきメールがなくなるまで、S1～S6の処理を繰り返す。そして、S1の処理によって確認した結果、メールサーバS₅に受信すべきメールがなくなると（S1:No）、このメール受信処理を終了する。本実施例の電子メールフィルタリングプログラムに従う上記のメール受信処理の実行によって、不特定多数の送信側端末Tsから送信された電子メールがスパムメールであるか否かを判定することができる。

30

【0098】

次に、図5のフローチャートを参照して、上記した経路情報取得処理（S3）について説明する。図5は、経路情報取得処理（S3）を示すフローチャートである。図5に示すように、経路情報取得処理（S3）は、まず、受信メールメモリ14aに記憶されているメールの「Received:」フィールドを参照し、配信経路上の中継装置（メールサーバS及び送信側端末Ts）のIPアドレスを取得し（S301）、取得したIPアドレスを経路情報メモリ14bに記憶し（S302）、この経路情報取得処理（S3）を終了する。上記した経路情報取得処理（S3）によって、配信経路上の送信側端末Ts及びメールサーバSを特定することができる。

40

【0099】

次に、図6のフローチャートを参照して、上記した配信木構築処理（S4）について説明する。図6は、配信木構築処理（S4）を示すフローチャートである。なお、「配信木」とは、配信経路を木構造で表したものである。この配信木は、「根」である受信側端末Trと、「ノード（所謂「節」とみなされる）」であるメールサーバSと、「葉」である送信側端末Tsとから構成される。

【0100】

図6に示すように、配信木構築処理（S4）は、まず、経路情報メモリ14bに記憶さ

50

れているIPアドレスの中から最も受信側端末に近い中継装置であるメールサーバS（図1ではS₅）のIPアドレスを読み込み（S401）、読み込んだIPアドレスが中継装置メモリ16a（第1中継装置メモリ16a₁～第n中継装置メモリ16a_n）に既存する中継装置のIPアドレスであるかを確認する（S402）。

【0101】

S402の処理により確認した結果、読み込んだIPアドレスが中継装置メモリ16aに既存する中継装置のものであれば（S402：Yes）、対応するスパム中継カウンタ16x₁及びハム中継カウンタ16x₂（xは1～nのうち対応する値）に記憶されている値を、その中継装置の配信木情報として配信木情報メモリ14cに記憶する（S403）。

10

【0102】

一方で、S402の処理により確認した結果、既存する中継装置のIPアドレスでなければ（S402：No）、新規の配信木情報として配信木情報メモリ14cに記憶する（S405）。

【0103】

S403又はS405の処理後、経路情報メモリ14bから読み込んだIPアドレスが送信元のIPアドレス、即ち、送信側端末T_sのIPアドレスであるか否かを確認し（S404）、そうでなければ（S404：No）、経路情報メモリ14bに記憶されているIPアドレスの中で、直前に読み込んだIPアドレスに対する1段上位（送信側端末T_sの側）のIPアドレスを読み込み（S406）、S402の処理へ移行する。そして、S404において、読み込んだIPアドレスが送信元のIPアドレスであることが確認されるまで、S402～S406を繰り返す。

20

【0104】

S404の処理により確認した結果、読み込んだIPアドレスが送信元のIPアドレスであれば（S404：Yes）、この配信木構築処理（S4）を終了する。この配信木構築処理（S4）により、1の配信経路を表す木構造である配信木が構築される。

【0105】

次に、図7のフローチャートを参照して、上記したメール判定処理（S5）について説明する。図7は、メール判定処理（S5）を示すフローチャートである。図7に示すように、メール判定処理（S5）では、まず、配信木情報メモリ14cに記憶されている1の配信木情報を読み出し（S501）、その配信木情報が新規の中継装置に対するものであるかを確認する（S502）。

30

【0106】

S502の処理により確認した結果、配信木情報が新規の中継装置に対するものでなければ（S502：No）、その中継装置に対する配信木情報（その中継装置が過去に中継したスパムメールの数及びハムメールの数）とスパム受信カウンタ16b及びハム受信カウンタ16cの値とを用いて、即ち、その中継装置に対するメール情報を用いて、下記式（1）に従ってベイズ確率モデルに基づくベイズ確率「pg_n」を求める（S503）。

【0107】

【数1】

40

$$pg_n = \frac{\frac{b}{nbad}}{\frac{g}{ngood} + \frac{b}{nbad}} \quad (1)$$

式（1）において、「b」は、その中継装置が過去に中継したスパムメールの数、即ち、その中継装置に対応するスパム中継カウンタ16a_x₁（x=1～n）の値であり、「g」は、その中継装置が過去に中継したハムメールの数、即ち、その中継装置に対応する

50

ハム中継カウンタ $16a_{x2}$ ($x = 1 \sim n$) の値である。また、「nbad」は、スパム受信カウンタ $16b$ の値であり、「ngood」は、ハム受信カウンタ $16c$ の値である。なお、このベイズ確率 pg_n の上限は「0.99」であり、下限は「0.01」であるとする。

【0108】

一方で、S502の処理により確認した結果、配信木情報が新規の中継装置に対するものであれば(S502:Yes)、ベイズ確率 pg_n の値を「0.5」とする(S512)。このように、配信経路上に新規の中継装置が確認された場合には、ベイズ確率「 pg_n 」の値を所定の定数として処理するので、新たな配信経路が生じ易い初見メールに対しても対応可能となる。また、幅広い配信経路に対しても適用可能となる。

10

【0109】

S512又はS503の処理後、配信木情報メモリ $14c$ に記憶されている全ての配信木情報を読み出したかを確認し(S504)、まだ読み出していない配信木情報があれば(S504:No)、S501の処理へ戻り、全ての配信木情報が読み出されるまで、S501~S504、S512の処理を繰り返す。そして、S504において、全ての配信木情報が読み出されたと確認されたら(S504:Yes)、S505の処理へ移行する。

【0110】

S505では、S503又はS512の処理によって、配信経路上の各中継装置に対して得られたベイズ確率 pg_n の値から、配信経路全体のベイズ確率 pg を下記式(2)から求める(S505)。

20

【0111】

【数2】

$$pg = \frac{pg_1 \times pg_2 \times \dots \times pg_n}{pg_1 \times pg_2 \times \dots \times pg_n + (1 - pg_1) \times (1 - pg_2) \times \dots \times (1 - pg_n)} \quad (2)$$

30

S505の処理後、配信経路全体のベイズ確率 pg の値が0.9を超えるかを確認し(S506)、ベイズ確率 pg の値が0.9を超えていれば(S506:Yes)、受信メールメモリ $14a$ に記憶され現在判定中のメールをスパムメールであると判定する(S513)。

【0112】

一方で、S506の処理により確認した結果、ベイズ確率 pg の値が0.9以下であれば(S506:No)、ベイズ確率 pg の値が0.1より小さいかを確認し(S507)、ベイズ確率 pg の値が0.1より小さければ(S507:Yes)、受信メールメモリ $14a$ に記憶され現在判定中のメールをハムメールであると判定する(S514)。

【0113】

また、S507の処理により確認した結果、ベイズ確率 pg の値が0.1以上であれば、即ち、ベイズ確率 pg が0.1以上かつ0.9以下である場合には(S507:No)、受信メールメモリ $14a$ に記憶され現在判定中のメールを、スパムメールともハムメールとも判定がつかず不確定要素の高いメールであるグレイメール(特許請求の範囲における「不確定メール」に該当する)であると判定する(S508)。

40

【0114】

S508の処理後、スパムメールともハムメールとも判定できないとされたグレイメールに対し、別のフィルタをかけてスパムメールであるかハムメールであるかを分類するグレイメール再判定処理を実行する(S509)。このグレイメール再判定処理(S509)では、例えば、グレイメールの本文中のテキスト情報に対して従来のテキストフィルタ

50

を適用することによって、グレイメールをスパムメール又はハムメールのいずれかに分類することができる。

【0115】

S509、S513又はS514の処理後、ベイズ確率 p_g に基づいてスパムメールと判定されたメール、及びS509におけるグレイメールの再判定によってスパムメールと判定されたメールに対して行う処理であるスパムメール対応処理を実行し(S510)、配信木情報メモリ14c及びメール受信メモリ14aをクリアし(S511)、このメール判定処理(S5)を終了する。

【0116】

なお、上記のスパムメール対応処理(S510)として実行できる処理としては、例えば、スパムメールと判定されたメールを受信メールメモリ14aから削除したり、そのスパムメールを受信拒否としたり、そのスパムメールを専用のフォルダに格納したり、スパムメールが受信されたことをユーザに報知したりなどの各種処理が挙げられる。このように、スパムメールと判定されたメールに対して削除や受信拒否するなどのスパムメール対応処理(S510)を実行することによって、ユーザがスパムメールにより受ける実害を低減することができる。なお、スパムメール対応処理(S510)においてスパムメールに対して所定の処理を行う一方で、ハムメールと判定されたメールについては、受信メールメモリ14aから記憶部16に設けられたメール格納部(非図示)に不揮発的に記憶するように構成すればよい。

【0117】

また、このメール判定処理(S5)におけるS508、S513又はS514の処理の結果として、受信メールメモリ14aに一時的に記憶されているメールが、スパムメール、ハムメール、グレイメールの3種類に分類される。ここで、図8を参照して、グレイメールについて説明する。

【0118】

図8は、図1に示した配信経路におけるスパムメール及びハムメールの偏りを説明するための模式図である。なお、図8において、丸括弧内に記載された数値は、左側が「その中継サーバを経由したハムメールの数(ハムメールの経由数)」を示す数値であり、右側が「その中継サーバを経由したスパムメールの数(スパムメールの経由数)」を示す数値である。即ち、左側の数値が、各中継サーバに対するハム中継カウンタ 16_{x2} ($x=1 \sim n$)の値であり、右側の数値が、各中継サーバに対するスパム中継カウンタ 16_{x1} ($x=1 \sim n$)の値である。

【0119】

図8に示すように、送信側端末 Ts_1 及び Ts_2 からスパムメールのみが送られた場合、これらの2つの配信経路における重なり部分であるメールサーバ S_1 、 S_4 、 S_5 を含む経路Aを通るメールは、スパムメールである可能性が高い。その結果として、メール判定処理(S5)では、経路Aを通るメールをスパムメールとして判定する。

【0120】

一方で、送信側端末 Ts_5 及び Ts_6 からハムメールのみが送られた場合、これらの2つの配信経路における重なり部分であるメールサーバ S_3 、 S_4 、 S_5 を含む経路Cを通るメールは、初見メールであってもハムメールである可能性が高い。その結果として、メール判定処理(S5)では、経路Cを経由するメールをハムメールとして判定する。

【0121】

しかし、メールサーバ S_2 、 S_4 、 S_5 を含む経路Bのように、その経路Bを経由するメールが、スパムメール又はハムメールのいずれでもあり得る場合には、その経路Bを経由するメールがスパムメールかハムメールかを区別し難い。ここで、そのように区別し難いメールに対し、スパムメールであるかハムメールであるかを厳密に区別を付けるような判定を行った場合には、それが誤判定となる可能性が高い。

【0122】

上記したメール判定処理(S5)では、スパムメールであるかハムメールであるかを区

10

20

30

40

50

別し難いメール、即ち、バイズ確率 p_g が 0.1 以上かつ 0.9 以下であったメールはグレイメールとして判定されるので、そのような誤判定を防止することができる。

【0123】

さらに、グレイメール再判定処理 (S509) において、グレイメールと判定されたメールに対して別のフィルタをかけることによって、スパムメールであるかハムメールであるかの判定精度を向上させることができる。

【0124】

次に、図9のフローチャートを参照して、上記した学習処理 (S6) について説明する。図9は、学習処理 (S6) を示すフローチャートである。図9に示すように、学習処理 (S6) は、まず、経路情報メモリ14bに記憶されているIPアドレスの中から最も受信側端末に近い中継装置であるメールサーバS (図1ではS₅) のIPアドレスを読み込み (S601)、読み込んだIPアドレスが中継装置メモリ16a (第1中継装置メモリ16a₁ ~ 第n中継装置メモリ16a_n) に既存する中継装置のIPアドレスであるかを確認する (S602)。

10

【0125】

S602の処理により確認した結果、読み込んだIPアドレスが中継装置メモリ16aに既存する中継装置のものであれば (S602: Yes)、対応するスパム中継カウンタ16x₁ 及びハム中継カウンタ16x₂ (xは1~nのうち対応する値) に記憶されている値を、S513又はS514で判定されたメールの判定結果に基づいて更新する (S603)。

20

【0126】

一方で、S602の処理により確認した結果、読み込んだIPアドレスが中継装置メモリ16aに既存する中継装置のIPアドレスでなければ (S602: No)、そのIPアドレスに対応する中継装置を新規の中継装置として中継装置メモリ16aに記憶する (S606)。即ち、その新たな中継装置に対応する新たな第n中継装置メモリ16a_n を中継装置メモリ16aに作成し、スパム中継カウンタ16a_{n1} 及びハム中継カウンタ16a_{n2} の値を、S513又はS514で判定されたメールの判定結果に基づいて「0」又は「1」に設定する。

【0127】

S606又はS603の処理後、スパム受信カウンタ16b又はハム受信カウンタ16cを、S513又はS514で判定されたメールの判定結果に基づいて更新する (S604)。次いで、経路情報メモリ14bから読み込んだIPアドレスが送信元のIPアドレス、即ち、送信側端末T_sのIPアドレスであるか否かを確認し (S605)、そうでなければ (S605: No)、経路情報メモリ14bに記憶されているIPアドレスの中で、直前に読み込んだIPアドレスに対する1段上位 (送信側端末T_sの側) のIPアドレスを読み込み (S607)、S602の処理へ移行する。そして、S605において、読み込んだIPアドレスが送信元のIPアドレスであることが確認されるまで、S602~S607を繰り返す。S605の処理により確認した結果、読み込んだIPアドレスが送信元のIPアドレスであれば (S605: Yes)、経路情報メモリ14bをクリアし (S608)、この学習処理 (S6) を終了する。

30

40

【0128】

この学習処理 (S6) によって、メール判定処理 (S5) によるメールに対する判定結果に応じて、記憶部16 (中継装置メモリ16a、スパム受信カウンタ16b、ハム受信カウンタ16c) に記憶されるメール情報が更新される。即ち、メール情報がメールの判定結果に応じて学習されていくので、学習すればするほどその判定精度を向上させることができる。

【0129】

上記のように機能する第1実施例の電子メールフィルタリングシステムによるメールの判定能力を検証した。なお、第1実施例の電子メールフィルタリングシステムを機能させるための電子メールフィルタリングプログラムは、C言語を用いてUNIX (登録商標)

50

上に作成した。

【 0 1 3 0 】

この検証において3種類のサンプルセットを用いた。この3種類のサンプルセットとは、表1に示す「サンプルセット1」、表2に示す「サンプルセット2」、表3に示す「サンプルセット3」である。なお、検証に使用されるすべての電子メールは、複数のメールサーバにより実際に受信されたものである。

【 0 1 3 1 】

【表1】

	テストメール		学習メール	
	ハムメール	スパムメール	ハムメール	スパムメール
日本語	198	1	1500	0
英語	2	199	0	1500

10

【 0 1 3 2 】

【表2】

	テストメール		学習メール	
	ハムメール	スパムメール	ハムメール	スパムメール
日本語	200	200	300	300
英語	0	0	0	0

20

【 0 1 3 3 】

【表3】

	テストメール		学習メール	
	ハムメール	スパムメール	ハムメール	スパムメール
日本語	200	100	800	400
英語	0	100	0	400

30

表1～表3に示すように、この検証においてサンプルとするスパムメールは、サンプルセット1では英語で記述されているものとし、サンプルセット2では日本語で記述されているものとし、サンプルセット3では半分が日本語で記述されたものであり、残りの半分が英語で記述されたものとした。

【 0 1 3 4 】

なお、表1～表3に示す各サンプルセットにおいて、「学習メール」とは、本実施例の電子メールフィルタリングシステムにおいて判定されたメールの判定結果を学習させるメールであり、「テストメール」とは、学習による判定精度向上の効果を比較するために、メールの判定結果を学習させないメールである。

【 0 1 3 5 】

上記のサンプルセット1～3に対し、第1実施例の電子メールフィルタリングシステムを適用することによってメールの判定を行った。即ち、メールの「Received:」フィールドから取得したIPアドレスのみを使用したメール判定を行った。その結果を表4～表6に示す。なお、以下において、「ham」及び「spam」の表記は、それぞれ、ハムメール及びスパムメールを示す。

40

【 0 1 3 6 】

【表 4】

学習数	ham を ham	ham を spam	spam を spam	spam を ham
10	39	0	0	0
20	76	0	0	0
30	76	0	0	0
40	77	0	0	0
50	93	0	0	0
60	97	0	0	0
70	104	0	0	0
80	104	0	0	0
90	104	0	0	0
100	112	0	0	0
200	125	2	0	0
300	135	1	0	0
400	141	1	0	0
500	142	0	0	0
1000	165	1	0	0
1500	170	0	1	0

10

【 0 1 3 7 】

【表 5】

学習数	ham を ham	ham を spam	spam を spam	spam を ham
10	39	0	0	0
20	76	0	0	0
30	76	0	2	0
40	77	0	2	0
50	93	0	2	0
60	97	0	4	0
70	104	0	4	0
80	104	0	4	0
90	104	1	4	0
100	112	1	4	0
200	125	1	8	0
300	136	1	14	0

20

【 0 1 3 8 】

【表 6】

学習数	ham を ham	ham を spam	spam を spam	spam を ham
10	39	0	0	0
20	76	0	0	0
30	76	0	2	0
40	77	0	2	0
50	93	0	2	0
60	97	0	2	0
70	104	0	2	0
80	104	0	2	0
90	104	0	2	0
100	112	0	2	0
200	125	0	2	0
300	136	1	2	0
400	141	2	8	0
500	142	1	8	0
800	163	1	8	0

30

40

表 4 ~ 表 6 に示すように、サンプルセット 1 ~ 3 のいずれも、学習メールが多ければ多いほどハムメールを正しく判定した。即ち、メールの言語に依存することなく、ハムメールを正しく判定した。また、スパムメールをハムメールとして誤判定されることはなかった。

50

【 0 1 3 9 】

本実施例の電子メールフィルタリングシステムでは、学習用のハムメールのIPアドレスと重複しない限り、スパムメールがハムメールとして誤判定されない。仮に、「Received:」フィールドに複数のIPアドレスを偽装して付加したとしても、新規の中継装置のIPアドレスはベイズ確率 p_{gn} が0.5として処理されるので、ハムメールと誤判定されることがないのである。

【 0 1 4 0 】

次に、図10及び図11を参照して、本発明の電子メールフィルタリングシステムの第2実施例について説明する。なお、この第2実施例において、上記した第1実施例と同一の部分には同一の符号を付して、その説明を省略する。

10

【 0 1 4 1 】

上記した第1実施例の電子メールフィルタリングシステムでは、メールの「Received:」フィールドから取得したIPアドレスのみから特定されたメールサーバに基づいてメールの判定を行った。しかし、不特定多数の送信側端末Tsから1の受信側端末Trまでの各配信経路において、配信経路に重複部分が少ない場合には、初見メールが「スパムメールが中継される傾向にある経路」を通過するか、「ハムメールが中継される傾向にある経路」を通過するかを判定するための情報が特に少なくなる。その場合、初見メールの判定精度に影響が生じることになる。

【 0 1 4 2 】

そこで、この第2実施例の電子メールフィルタリングシステムでは、第1実施例における経路情報取得処理(S3)におけるS302の処理後、中継装置として推定されるメールサーバを配信経路上に補完するための経路追跡処理(S303)を実行する。なお、以下の説明では、経路追跡処理(S303)によって推定されて補完されたメールサーバを便宜的にメールサーバ'として表すことがある。

20

【 0 1 4 3 】

ここで、図10は、第2実施例の電子メールフィルタリングシステムにおける経路情報取得処理(S3)のフローチャートである。図10に示すように、経路追跡処理(S303)は、以下に説明するS303a~S303dの各ステップから構成される。

【 0 1 4 4 】

まず、S303aの処理として、中継ルータの探索を行い、そのIPアドレスを取得する。このS303aの処理において行う中継ルータの探索は、Traceroute(例えば、W. Richard Stevens、「詳解TCP/IP Vol.1 プロトコル」Pearson Education Japan, 2004年6月20日新装版第5刷を参照のこと)などの技術を使用することができる。Tracerouteの実行により、目的のホストまでに経由した中継ルータのIPアドレスを調べることができる。なお、本実施例では、このTracerouteを実行する上で、「Received:」フィールドに記録されているIPアドレス群のうち、最も受信者に近いIPアドレスを使用する。

30

【 0 1 4 5 】

S303aの処理後、取得した中継ルータのIPアドレスが所属するネームサーバの情報(ネームサーバ情報)を取得する(S303b)。このS303bの処理において行うネームサーバ情報の取得は、例えば、Whoisプロトコルを使用するWhoisサーバへの問い合わせを利用することができる。なお、Whoisプロトコルとは、ドメイン名の登録やIPアドレスの割当に関する情報をオンラインで提供する仕組みである。このWhoisサーバへの問い合わせにより取得されるネームサーバ情報は、ネームサーバのドメイン名である。

40

【 0 1 4 6 】

S303bの処理後、取得したネームサーバ情報によって表されるネームサーバにDNS(Domain Name System)の問い合わせを行い、メールの配信経路上において補完されるメールサーバS'のIPアドレスを取得し(S303c)。取得したメールサーバS'のIPアドレスを経路情報メモリ14bに記憶し(S303d)、第2実施例の経路情報取得処理(S3)を終了する。

【 0 1 4 7 】

50

DNSサーバには、Aレコード(名前 IPアドレスの定義)やPTRレコード(IPアドレス 名前の定義)、NSレコード(ネームサーバの定義)、SOAレコード(ドメインのオーソリティ情報の定義)、MXレコードなど、さまざまなレコード(情報)が登録されている。ここで、MXレコードは、Mail Exchangerの略であり、そのドメインにおけるメールサーバに関する情報が登録されている。

【0148】

あるドメインから別のドメインに対してメール送信しようとするとき、送信元のメールサーバは、送信先ドメインのDNSサーバに対してMXレコードの情報を問い合わせる。MXレコードには、当該ドメインにおけるメールサーバとなっているコンピュータの名前(FQDN名)とIPアドレスなどの情報が含まれており、これを基にしてメールの送信先(通常はSMTPのサーバ)を知ることができるのである。

10

【0149】

S303cの処理では、DNSに対してMXレコードへの参照を要求する。このアクセスは、リゾルバ(resolver)を経由して行われる。例えば、「xxx.ac.jp」のMXレコードの情報を問い合わせた結果が以下の通りであれば、補完されるメールサーバS'のIPアドレスは、「133.15.xxx.1」である。

【0150】

```
xxx.xx.jp MX preference = 0, mail exchanger = server1.xxx.xx.jp
xxx.xx.jp MX preference = 100, mail exchanger = server.xxx.xx.jp
xxx.xx.jp nameserver = server.xxx.xx.jp
xxx.xx.jp nameserver = nameserv.gw.xxx-u.xx.jp
server1.xxx.xx.jp      internet address = 133.15.xxx.1
server.xxx.xx.jp      internet address = 133.15.xxx.1
nameserv.gw.xxx-u.xx.jp internet address = 192.50.xx.9
```

20

【0151】

ここで、図11を参照して、上記の経路情報取得処理(S3)における経路追跡処理(S303)による結果を模式的に説明する。図11(a)は、各配信経路に重複部分がない場合を模式的に示す図であり、図11(b)は、経路追跡処理(S303)によってメールサーバS'(中継装置)が補完された状態を模式的に示す図である。

【0152】

上記したS303a~S303dから構成される経路追跡処理(S303)が実行された結果、メールの配信経路は図11(b)に示す通りとなる。即ち、図11(a)に示した配信経路における、送信側端末Tsに接続されるメールサーバSと受信側端末に接続されるメールサーバSとの間にメールサーバS'が補完される。配信経路上にメールサーバS'を補完することによって、各配信経路上において重複する中継装置を増やすことができる。その結果として、スパムメールが中継され易い経路と、ハムメールが中継され易い経路とを確率的に区別し易くすることができるのである。

30

【0153】

S303の処理後、取得したメールサーバS'のIPアドレスを経路情報メモリ14bに格納し(S303d)、この経路情報取得処理(S3)を終了する。なお、この第2実施例では、配信木における「ノード」には、メールサーバSとして補完されたメールサーバS'も含むものとする。

40

【0154】

サンプルセット1(表1参照)に対し、中継ルータの探索(経路追跡)を行いメールサーバS'の補完を行った場合と、行わない場合とについて、配信経路上におけるメールサーバS(補完されたメールサーバS'を含む)の数の変化を表7に示す。

【0155】

【表 7】

深さ	1	2	3	4	5	6	7	8	9	10	平均
ham(経路追跡前)	89	72	11	7	0	20	1	0	0	0	2.1
spam(経路追跡前)	191	8	1	0	0	0	0	0	0	0	1.1
ham(経路追跡後)	11	27	34	26	42	22	28	9	1	0	4.5
spam(経路追跡後)	0	0	1	90	56	31	16	4	1	1	5.0

なお、表 7 において、「経路追跡前」とは、メールサーバ S' の補完を行う前、即ち、メールの「Received:」フィールドから取得した IP アドレスのみからメールサーバ S を特定した場合を示す。一方、「経路追跡後」とは、上記した経路追跡処理 (S 3 0 3) によってメールサーバ S' が補完された場合を示す。

【 0 1 5 6 】

表 7 に示すように、ハムメール及びスパムメールのいずれも、経路追跡前の状態では、根 (受信側端末 Tr) から葉 (送信側端末 Ts) までの深さの大半が 1 または 2 であった。この状態では、配信経路が重複することは少ない。

【 0 1 5 7 】

一方で、経路追跡後、即ち、経路追跡処理 (S 3 0 3) の実行後は、両メールとも、同一の経路を経由する重なりが増え、深さの平均が 4 を超えた。よって、経路追跡処理 (S 3 0 3) の実行によって、よりハムメール的な経路と、よりスパムメール的な経路と、グレイメール的な経路との区別をより明確にすることができ、その結果、スパムメールであるか又はハムメールであるかの判定精度をより向上させることができるのである。

【 0 1 5 8 】

次に、第 2 実施例の電子メールフィルタリングシステム及び従来のホワイトリストフィルタについて、ハムメールの判定精度を比較した。図 1 2 は、第 2 実施例の電子メールフィルタリングシステムによるハムメールに対する判定結果と従来のホワイトリストフィルタによるハムメールに対する判定結果とを記したグラフである。なお、ホワイトリストフィルタに用いるホワイトリストは、サンプルセット 1 ~ 3 のそれぞれにおける「学習メール」としたメールのメールアドレスを全て登録することによって作成した。

【 0 1 5 9 】

図 1 2 (a) ~ (c) は、それぞれサンプルセット 1 ~ 3 に対する結果を示すグラフである。ここで、グラフ 1 2 0 a , 1 2 1 a , 1 2 2 a は、それぞれ、サンプルセット 1 ~ 3 について、本実施例の電子メールフィルタリングシステムによってハムメールを正しく判定した検出率を示すグラフである。一方で、グラフ 1 2 0 b , 1 2 1 b , 1 2 2 b は、それぞれ、サンプルセット 1 ~ 3 について、従来のホワイトリストフィルタによってハムメールを正しく判定した検出率を示すグラフである。また、グラフ 1 2 0 c , 1 2 1 c , 1 2 2 c は、それぞれ、サンプルセット 1 ~ 3 について、本実施例の電子メールフィルタリングシステムによりハムメールをスパムメールとして誤判定した検出率を示すグラフである。一方で、グラフ 1 2 0 d , 1 2 1 d , 1 2 2 d は、それぞれ、サンプルセット 1 ~ 3 について、従来のホワイトリストフィルタによりハムメールをスパムメールとして誤判定した検出率を示すグラフである。

【 0 1 6 0 】

図 1 2 (a) ~ 図 1 2 (c) に示すように、サンプルセット 1 ~ 3 のいずれの場合も、本実施例の電子メールフィルタリングシステムを用いた方が、従来のホワイトリストフィルタを用いた場合に比べ、ハムメールの判定精度が全体に渡って良好であった。ここで、図 1 2 (a) に示すように、サンプルセット 1 に対しては、学習メールが 1 0 0 通及び 1 0 0 0 通の場合に、ハムメールの判定精度がそれぞれ 6 2 . 0 % 及び 1 6 . 5 % 向上した。また、図 1 2 (b) に示すように、サンプルセット 2 に対しては、学習メールが 1 0 0 通及び 3 0 0 通の場合に、ハムメールの判定精度がそれぞれ 6 2 . 0 % 及び 2 8 . 5 % 向上した。また、図 1 2 (c) に示すように、サンプルセット 3 に対しては、学習メールが

10

20

30

40

50

100通及び500通の場合に、ハムメールの判定精度がそれぞれ60.5%及び25.5%向上した。この結果は、従来のホワイトフィルタでは初見メールに対する判定が困難であるのに対し、本実施例の電子メールフィルタリングシステムでは、配信経路上の中継装置におけるメール情報に基づいてメールを判定するので、初見メールであっても精度よくハムメールを判定できることを示している。また、日本語、英語といった言語情報の差違にかかわらず、精度よくハムメールを判定できることを示している。

【0161】

次に、第2実施例の電子メールフィルタリングシステム及び従来のブラックリストフィルタについて、スパムメールの判定精度を比較した。図13は、第2実施例の電子メールフィルタリングシステムによるスパムメールに対する判定結果と従来のブラックリストフィルタによるスパムメールに対する判定結果とを記したグラフである。なお、ブラックリストへの登録内容は、学習メールにおけるスパムメールのメールアドレスを全て登録した。同時に、DNSブラックリストに登録されている20万件のIPアドレスを使用した。

10

【0162】

図13(a)~(c)は、それぞれサンプルセット1~3に対する結果を示すグラフである。ここで、グラフ130a, 131a, 132aは、それぞれ、サンプルセット1~3について、本実施例の電子メールフィルタリングシステムによってスパムメールを正しく判定した検出率を示すグラフである。一方で、グラフ130b, 131b, 132bは、それぞれ、サンプルセット1~3について、従来のブラックリストフィルタによってスパムメールを正しく判定した検出率を示すグラフである。また、グラフ130c, 131c, 132cは、それぞれ、サンプルセット1~3について、本実施例の電子メールフィルタリングシステムによりスパムメールをハムメールとして誤判定した検出率を示すグラフである。一方で、グラフ130d, 131d, 132dは、それぞれ、サンプルセット1~3について、従来のブラックリストフィルタによりスパムメールをハムメールとして誤判定した検出率を示すグラフである。

20

【0163】

図13(a)~図13(c)に示すように、サンプルセット1~3のいずれの場合も、本実施例の電子メールフィルタリングシステムを用いた方が、従来のブラックリストフィルタを用いた場合に比べ、スパムメールの判定精度が全体に渡って良好であった。ここで、図13(a)に示すように、サンプルセット1に対しては、学習メールが100通及び1000通の場合にいずれも、スパムメールの判定精度が89.0%向上した。また、図13(b)に示すように、サンプルセット2に対しては、学習メールが100通及び300通の場合に、スパムメールの判定精度がそれぞれ16.0%及び25.5%向上した。また、図13(c)に示すように、サンプルセット3に対しては、学習メールが100通及び500通の場合に、スパムメールの判定精度がそれぞれ60.5%及び25.5%向上した。

30

【0164】

よって、本実施例の電子メールフィルタリングシステムによれば、サンプルセット1のような送信元の地域を日本と海外とである程度の分類が可能である場合に、ブラックリストフィルタよりスパムメールの判定精度が格段に向上することを示す。また、サンプルセット2のような、発信元が主に日本である故にハムメール及びスパムメールの発信元地域を分類し難い場合であっても、学習が行われることによって判定精度が向上することを示している。また、日本語、英語といった言語情報の差違にかかわらず、精度よくスパムメールを判定できることを示している。

40

【0165】

次に、図14を参照して、第2実施例の電子メールフィルタリングシステムにおけるグレイメール再判定処理(S509:図7参照)の実行による効果について検証する。図14は、サンプルセット3について、グレイメール再判定処理(S509)を実行した場合と実行しなかった場合におけるメールの判定精度を比較するグラフである。図14(a)は、ハムメールに対する判定精度を示すグラフであり、図14(b)は、スパムメールに

50

対する判定精度を示すグラフである。なお、テキスト型ベイジアンフィルタとして、「bs filter」(<http://bsfilter.org/>)を用い、Paul Graham方式で実行した。

【0166】

図14(a)において、グラフ140aは、グレイメール再判定処理(S509)を実行した場合におけるハムメールを正しく判定した検出率を示すグラフである。一方で、グラフ140bは、グレイメール再判定処理(S509)を実行しなかった場合におけるハムメールを正しく判定した検出率を示すグラフである。また、グラフ140cは、グレイメール再判定処理(S509)を実行した場合に、スパムメールをハムメールとして誤判定した検出率を示すグラフである。一方で、グラフ140dは、グレイメール再判定処理(S509)を実行しなかった場合に、スパムメールをハムメールとして誤判定した検出率を示すグラフである。

10

【0167】

図14(b)において、グラフ141aは、グレイメール再判定処理(S509)を実行した場合におけるスパムメールを正しく判定した検出率を示すグラフである。一方で、グラフ141bは、グレイメール再判定処理(S509)を実行しなかった場合におけるスパムメールを正しく判定した検出率を示すグラフである。また、グラフ141cは、グレイメール再判定処理(S509)を実行した場合に、ハムメールをスパムメールとして誤判定した検出率を示すグラフである。一方で、グラフ141dは、グレイメール再判定処理(S509)を実行しなかった場合に、ハムメールをスパムメールとして誤判定した検出率を示すグラフである。

20

【0168】

図14(a)に示すように、ハムメールの判定精度は、学習メールが100通及び500通の場合にそれぞれ13.5%及び5.5%向上した。一方で、スパムメールをハムメールとして誤判定した検出率が悪化した(学習メール100通の場合に29.0%、学習メール500通の場合に15.5%)が、これは、グレイメール再判定処理(S509)を実行した結果として生じたハムメールの誤判定率の悪化(グラフ140dにおける各値に対するグラフ140cにおける各値の上昇)は、ハムメールでもスパムメールでもない第3の区分であるグレイメールを考慮した結果として、グレイメールに対する再判定を実行しなかった場合における誤判定率の低下分がグレイメールに振り分けられたことに起因すると考えられる。

30

【0169】

また、図14(b)に示すように、学習メールが100通及び500通の場合に、スパムメールの判定精度がそれぞれ8.5%及び17.5%向上した。一方で、ハムメールをスパムメールとして誤判定した検出率には変化がなかった。この結果は、グレイメールに対しテキスト型ベイジアンフィルタを適用して再判定を行うことによって、ハムメール及びスパムメールの判定精度が向上することを示している。

【0170】

次に、グレイメール再判定処理(S509:図7参照)としてテキスト型ベイジアンフィルタを併用してメール判定を行った場合と、第2実施例の電子メールフィルタリングシステムに換えてテキスト型ベイジアンフィルタのみを用いてメール判定を行った場合について比較した。

40

【0171】

図15は、サンプルセット3について、本実施例の電子メールフィルタリングシステムにベイジアンフィルタを併用した場合の判定結果と、テキスト型ベイジアンフィルタのみを使用した場合の判定結果とを比較するグラフである。図15(a)は、ハムメールに対する判定精度を示すグラフであり、図15(b)は、スパムメールに対する判定精度を示すグラフである。なお、テキスト型ベイジアンフィルタは、図14に示した結果を得るために用いたものと同じく「bsfilter」を用い、Paul Graham方式で実行した。

【0172】

図15(a)において、グラフ150aは、本実施例の電子メールフィルタリングシス

50

テムにテキスト型ベイジアンフィルタを併用した場合におけるハムメールを正しく判定した検出率を示すグラフである。一方で、グラフ150bは、本実施例の電子メールフィルタリングシステムに換えてテキスト型ベイジアンフィルタを使用した場合におけるハムメールを正しく判定した検出率を示すグラフである。また、グラフ150cは、本実施例の電子メールフィルタリングシステムにテキスト型ベイジアンフィルタを併用した場合に、スパムメールをハムメールとして誤判定した検出率を示すグラフである。一方で、グラフ150dは、本実施例の電子メールフィルタリングシステムに換えてテキスト型ベイジアンフィルタを使用した場合に、スパムメールをハムメールとして誤判定した検出率を示すグラフである。

【0173】

図15(a)に示すように、学習メールが100通及び500通の場合に、ハムメールの誤判定精度をそれぞれ14.5%及び5.5%改善した。この結果は、本実施例の電子メールフィルタリングシステムにテキスト型ベイジアンフィルタを併用することによって、ハムメールの誤判定の程度を改善することを示している。即ち、従来のテキスト型ベイジアンフィルタでは、ハムメールに重みを置かれていることに基づく誤判定のし易さや、グレイメールを設けることなくハムメールかスパムメールかを二者択一的に判定することに基づく誤判定のし易さや、構文解析の難しい日本語に対する判定程度の低さなどが問題であったが、本実施例の電子メールフィルタリングシステムのようにテキスト型ベイジアンフィルタを併用することによってこれらの問題を解決し得ることを示している。

【0174】

ここで、この図15(a)におけるグラフ150cは、図14(a)におけるグラフ140cと同一のグラフである。よって、図14(a)と図15(a)とから、スパムメールをハムメールとして誤判定した検出率に関し、(1)グレイメールに対する再判定処理(S509:図7参照)を実行しない場合(グラフ140dに対応)と、(2)グレイメールに対し、従来のテキスト型ベイジアンフィルタを併用した場合(グラフ140c及びグラフ150cに対応)と、(3)従来のテキスト型ベイジアンフィルタを使用した場合(グラフ150dに対応する)とを比較することができる。

【0175】

よって、図14(a)に示すように、グレイメール再判定処理(S509)を実行した場合のハムメールの誤判定率(グラフ140c)が、グレイメール再判定処理(S509)を実行しなかった場合(グラフ140d)に比べて悪化したとしても、それは従来技術(従来のテキスト型ベイジアンフィルタ)に対する悪化を示すものではない。即ち、図15(a)によれば、グレイメール再判定処理(S509)を実行した結果として、ハムメールの誤判定率(グラフ150c(グラフ140cに対応))は、従来テキスト型ベイジアンフィルタにおけるハムメールの誤判定率(グラフ150d)に比べて改善されていることが明白である。

【0176】

一方、図15(b)において、グラフ151aは、本実施例の電子メールフィルタリングシステムにテキスト型ベイジアンフィルタを併用した場合におけるスパムメールを正しく判定した検出率を示すグラフである。一方で、グラフ151bは、本実施例の電子メールフィルタリングシステムに換えてテキスト型ベイジアンフィルタを使用した場合におけるスパムメールを正しく判定した検出率を示すグラフである。また、グラフ151cは、本実施例の電子メールフィルタリングシステムにテキスト型ベイジアンフィルタを併用した場合に、ハムメールをスパムメールとして誤判定した検出率を示すグラフである。一方で、グラフ151dは、本実施例の電子メールフィルタリングシステムに換えてテキスト型ベイジアンフィルタを使用した場合に、ハムメールをスパムメールとして誤判定した検出率を示すグラフである。

【0177】

図15(b)に示すように、学習メールが100通及び500通の場合に、スパムメールの判定精度がそれぞれ14.5%及び5.5%向上した。また、誤判定精度には変化が

10

20

30

40

50

なかった。この結果は、本実施例の電子メールフィルタリングシステムにテキスト型ページアンフィルタを併用することによって、スパムメールの判定精度が、テキスト型ページアンフィルタのみを使用する場合に比べて向上することを示している。

【0178】

次に、本発明の電子メールフィルタリングシステムの第3実施例について説明する。なお、この第3実施例において、上記した第1及び第2実施例と同一の部分には同一の符号を付して、その説明を省略する。

【0179】

この第3実施例の電子メールフィルタリングシステムは、「Received:」フィールドに記録されている情報が偽装された情報であることが確認された場合には、そのメールをスパムメールであると判定するものであり、「Received:」フィールドに記録されている情報が偽装されたものであるか否かを確認するために、IPアドレス確認処理(S7)を行う。

【0180】

このIPアドレス確認処理(S7)は、まず、メールの「Received:」フィールドに記録されているIPアドレスを、受信側に近い方から順にDNSサーバに問い合わせることによってFQDN名を逆引きする処理(S7a)と、その逆引きによって得られたFQDN名と「Received:」フィールドに記録されている送信ホスト名とが一致するかを確認する処理(S7b)とから構成される処理である。

【0181】

このIPアドレス確認処理(S7)は、その判定結果を学習させるために学習処理(S6)の前に実行することが好ましい。ここで、IPアドレス確認処理(S7)が、経路情報取得処理(S3)より前に実行される場合には、S7bの処理により確認した結果、逆引きによって得られたFQDN名と「Received:」フィールドに記録されている送信ホスト名とが不一致である場合には(S7b:No)、スパムメールであると判定し(S8)、その一方で、一致する場合には(S7b:Yes)、未判定メールと認定する(S9)。そして、未判定メールであった場合には、経路情報取得処理(S3)、配信木構築処理(S4)、メール判定処理(S5)を実行して、そのメールがスパムメールであるかを判定するように構成すればよい。なお、S8においてスパムメールと判定された場合には、そのメールがスパムメールであることを前提として経路情報取得処理(S3)、配信木構築処理(S4)、メール判定処理(S5)、学習処理(S6)を実行するように構成すればよい。

【0182】

あるいは、メール判定処理(S5)によりグレイメールと判定されたメールに対し、IPアドレス確認処理(S7)を実行するように構成してもよい。この場合には、メール判定処理(S5)におけるS508の処理以降、例えば、S508とS509との間などに実行することができる。即ち、S508の処理によりグレイメールと判定されたメールに対し、DNS逆引きによって得られたFQDN名と「Received:」フィールドに記録されている送信ホスト名とが不一致である場合には(S7b:No)、スパムメールであると判定し(S8)、その一方で、一致する場合には(S7b:Yes)、ハムメールと判定する(S10)ように構成すればよい。

【0183】

なお、このようにDNS逆引きによって得られたFQDN名と「Received:」フィールドに記録されている送信ホスト名とを比較する手法は、従来のメールサーバにセキュリティ機能として実装されている。しかし、従来のセキュリティ機能ではこの手法はホストの判定のみに使用されており、本実施例の電子メールフィルタリングシステムでは、一つのメールを中継する配信経路上のメールサーバSを一組とした情報として取り扱う。即ち、この手法をメールの配信経路の全体に対して適用し、メールが中継される経路からスパムメールかどうかを判定する。よって、中継装置が動的IPアドレス割り当てによって変化しても、柔軟にかつより詳細に対応できるのである。また、ヘッダが明らかに偽造されて

10

20

30

40

50

いるものをスパムメールとして処理できるので、処理を効率化できると共に、スパムメールであるか又はハムメールであるかの判定精度を向上させることができる。

【0184】

以上説明したように、本発明の電子メールフィルタリングシステムによれば、配信経路上の中継装置（送信側端末Ts及びメールサーバS）をIPアドレスによって特定した上で、その特定された中継装置が過去に中継した迷惑メール及び正当なメールの頻度を示すメール情報をメール判定のために利用することにより、配信経路上を通過して配信されるメールがスパムメールである場合にそれを確実に検出できると共に、正当なメールであるハムメールが過剰に拒否されることを抑制できる。

【0185】

また、配信経路上の中継装置（送信側端末Ts及びメールサーバS）が過去に中継した迷惑メール及び正当なメールの頻度を示すメール情報をメール判定のために利用するので、従来のテキストフィルタのような膨大なデータの蓄積を必要しない。そのため、データベースによる記憶装置（メモリやディスクなど）の消費量を抑制することができる。さらに、スパムメールであるか否かの判定を行う場合に、送信者による偽装が容易である電子メールに含まれるテキストを利用しないので、配信経路上を通過して配信される電子メールがスパムメールである場合にそれを確実に検出できると共に、正当なメールであるハムメールが過剰に拒否されることを抑制できる。

【0186】

さらに、上記のように、配信経路上の中継装置のIPアドレスを利用してメールの判定を行うので、言語情報に依存しない。よって、構文解析が難しくテキストフィルタでは誤判定されやすい日本語のメールに対しても、スパムメールであるかハムメールであるかを確実に検出することができる。

【0187】

加えて、迷惑メールであるか否かの判定を行う場合に、送信者による偽装が容易である送信者のアドレス（メールアドレス）を利用しないので、配信経路上を通過して配信されるメールがスパムメールである場合にそれを確実に検出できると共に、正当なメールであるハムメールが過剰に拒否されることを抑制できる。さらに、従来のアドレスフィルタのように、悪質な送信者との黜ごっこのような偽装アドレスの登録及び削除を繰り返す必要がなくなり、管理が容易になる。

【0188】

なお、請求項1記載の中継アドレス取得ステップ及び請求項7、13記載の中継アドレス取得手段としては、S301の処理が該当する。また、請求項1記載の迷惑メール中継確率取得ステップ及び請求項7、13記載の迷惑メール中継確率取得手段としては、S503及びS512の処理が該当する。また、請求項1記載の迷惑メール受信確率取得ステップ及び請求項7、13記載の迷惑メール受信確率取得手段としては、S505の処理が該当する。また、請求項1記載のメール判定ステップ及び請求項7、13記載のメール判定手段としては、S506～S508、S513、S514の処理及びS8の処理が該当する。また、請求項1記載の中継ルータアドレス取得ステップ及び請求項7、13記載の中継ルータアドレス取得手段としては、S303aの処理が該当する。また、請求項1記載の補完アドレス取得ステップ及び請求項7、13記載の補完アドレス取得手段としては、S303cの処理が該当する。

【0189】

また、請求項2記載の迷惑メール判定ステップ及び請求項8、14記載の迷惑メール判定手段としては、S506におけるYesの分岐処理及びS513の処理が該当する。また、請求項2記載の正当メール判定ステップ及び請求項8、14記載の正当メール判定手段としては、S507におけるYesの分岐処理及びS514の処理が該当する。また、請求項2記載の情報更新ステップ及び請求項8、14記載の情報更新手段としては、学習処理（S6）が該当する。また、請求項3記載のグレイメール認識ステップ及び請求項9、

10

20

30

40

50

15記載のグレイメール認識手段としては、S507におけるNoの分岐処理及びS508の処理が該当する。また、請求項4記載の不確定メール再判定ステップ及び請求項10、16記載の不確定メール再判定手段としては、グレイメール再判定処理(S509)が該当する。

【0190】

また、請求項5記載のアドレス確認ステップ及び請求項11、17記載のアドレス確認手段としては、IPアドレス確認処理(S7)が該当する。

10

【0191】

以上、実施例に基づき本発明を説明したが、本発明は上述した実施例に何ら限定されるものではなく、本発明の趣旨を逸脱しない範囲内で種々の改良変更が可能であることは容易に推察できるものである。

【0192】

例えば、上記実施例の電子メールフィルタリングシステムは、受信側端末Trに実装されるシステムであるとして説明したが、これに限定されず、メールサーバSや中継ルータに実装するシステムとして構成してもよい。また、一部の処理を受信側端末Trで行い、残りの処理をメールサーバSで行うなど、複数の装置で分割された処理が実行されて全体として電子メールフィルタリングシステムとして機能するように構成してもよい。

20

【0193】

また、上記実施例の電子メールフィルタリングシステムを機能させるための電子メールフィルタリングプログラムは、所謂パーソナルコンピュータである受信側端末TrのROM12に格納されている制御プログラム12aの一部であるとしたが、これに限定されるものではない。例えば、ファイアウォールやアプライアンスやエンタープライズなどの製品や装置に実装されて、これらの製品や装置において、上記実施例によって説明したような電子メールフィルタリングシステムを機能させるものであってもよい。また、格納場所はROMに限定されず、ハードディスクなどの書き換え可能なメモリや、各種記憶媒体(CD-ROMなど)などであってもよい。

【0194】

また、上記実施例の電子メールフィルタリングシステムでは、ベイズ確率 p_g の値が、0.1以上かつ0.9以下である場合に、そのメールをグレイメールであると判定するように構成したが、グレイメールとする範囲を設けることなく、スパムメールであるかハムメールであるかの二者択一の判定を行うように構成してもよい。例えば、ベイズ確率 p_g の値が、0.9以上であればスパムメールであると判定し、0.9より小さければハムメールであると判定するように構成してもよい。なお、スパムメールであるか、ハムメールであるか、グレイメールであるかを区分するための閾値は、必要に応じて適宜設定可能であることは容易に推察可能な事項である。また、本実施例では、S506において、ベイズ確率 $p_g > 0.9$ であるか否かを確認するように構成したが、ベイズ確率 $p_g = 0.9$ であるか否かを確認するように構成することも容易に推察可能な事項である。なお、S507の場合についても同様である。

30

40

【0195】

また、メールの判定区分をスパムメールであるかハムメールであるかの二者択一とする場合には、式1に換えて下記式(3)を用いるように構成してもよい。

【0196】

【数 3】

$$pg_n = \frac{\frac{b}{nbad}}{\frac{2 \times g}{ngood} + \frac{b}{nbad}} \quad (3)$$

10

また、この場合、本実施例においてベイズ確率「 pg_n 」に対して設定した上限「0.98」に換えて「0.99」とすることが好ましい。なお、ベイズ確率「 pg_n 」の値の上限及び下限についても必要に応じて適宜設定できることは容易に推察可能である。

【0197】

また、メールの判定区分をスパムメールであるかハムメールであるかの二者択一とする場合には、配信木情報が新規の中継装置に対するものであった場合に、ベイズ確率「 pg_n 」の値を「0.4」とすることが好ましいが、この定数の値もまた、適宜設定可能であることは容易に推察可能である。

【0198】

また、上記実施例では、グレイメール再判定処理（S509）において、グレイメールの本文中のテキスト情報に対して従来のテキストフィルタをかけるように説明したが、これに限定されるものではなく、グレイメールのメールヘッダにおけるメールアドレスなどの他の部分にテキストフィルタを適用するように構成してもよい。

20

【0199】

また、上記実施例では、S2において、1通ずつメールサーバSから受信するように構成したが、メールサーバSにあるメールを全て受信してから処理するように構成してもよい。また、上記実施例では、受信したメールをRAM14（受信メールメモリ14a）に一時的に記憶させるように構成したが、記憶部16に記憶させるように構成してもよい。

【0200】

また、上記実施例では、学習処理（S6）と配信木構築処理（S4）とをそれぞれ独立した処理として説明したが、学習処理（S6）におけるS603、S604、S606の処理を、配信木構築処理（S4）の中で実行されるように構成してもよい。

30

【0201】

また、上記実施例における配信木構築処理（S4）では、受信側端末Trに近い側の中継装置から順に、そのIPアドレスを経路情報メモリ14bから読み出すように構成したが、経路情報メモリ14bからIPアドレスを読み出す順序はこれに限定されるものではない。

【0202】

また、上記実施例では、受信したメールがメール判定手段（S5）によって判定された後に、その判定結果に基づいて学習処理（S6）が実行されるように構成されているが、学習処理（S6）は、ユーザが受信したメールに対してスパムメールであるか否かを判定した結果に基づいて実行するように構成してもよい。この場合は、図9に示した学習処理（S6）において、S601の処理後に、経路情報取得処理（S3）を実行して、ユーザが判定したメールの経路情報を取得するように構成すればよい。

40

【図面の簡単な説明】

【0203】

【図1】本発明の第1実施例における電子メールフィルタリングシステムが実装される電子メールの配信経路の一例を示す模式図である。

【図2】メールの「Received:」フィールドを示す模式図である。

【図3】本実施例の電子メールフィルタリングシステムを機能させる受信側端末の構成を

50

示すブロック図である。

【図4】受信側端末に実装された電子メールフィルタリングプログラムによって実行されるメール受信処理を示すフローチャートである。

【図5】経路情報取得処理を示すフローチャートである。

【図6】配信木構築処理を示すフローチャートである。

【図7】メール判定処理を示すフローチャートである。

【図8】図1の配信経路におけるメールの偏りを説明するための模式図である。

【図9】学習処理を示すフローチャートである。

【図10】第2実施例における経路情報取得処理を示すフローチャートである。

【図11】第2実施例の経路追跡処理による結果を説明するための模式図である。

10

【図12】第2実施例の電子メールフィルタリングシステムと従来のホワイトリストフィルタとを比較するグラフである。

【図13】第2実施例の電子メールフィルタリングシステムと従来のブラックリストフィルタとを比較するグラフである。

【図14】グレイメール再判定処理を実行した場合と実行しなかった場合におけるメールの判定精度を比較するグラフである。

【図15】テキスト型ベイジアンフィルタを併用した場合と、テキスト型ベイジアンフィルタのみを使用した場合とを比較するグラフである。

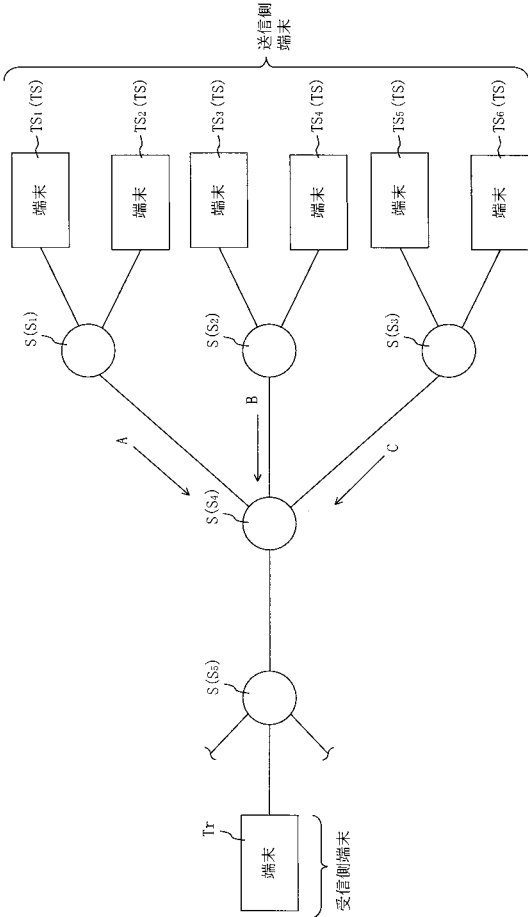
【符号の説明】

【0204】

20

16a (16a ₁ ~ 16a _n)	中継装置メモリ (情報記憶手段)
16a ₁₁ ~ 16a _{n1}	スパム中継カウンタ (情報記憶手段)
16a ₁₂ ~ 16a _{n2}	ハム中継カウンタ (情報記憶手段)
16b	スパム受信カウンタ (情報記憶手段)
16c	ハム受信カウンタ (情報記憶手段)
S	メールサーバ (中継装置)
Ts	送信側端末 (中継装置)
Tr	受信側端末 (コンピュータ)
R1, R2	「Received:」フィールド (ヘッダ情報)

【図1】

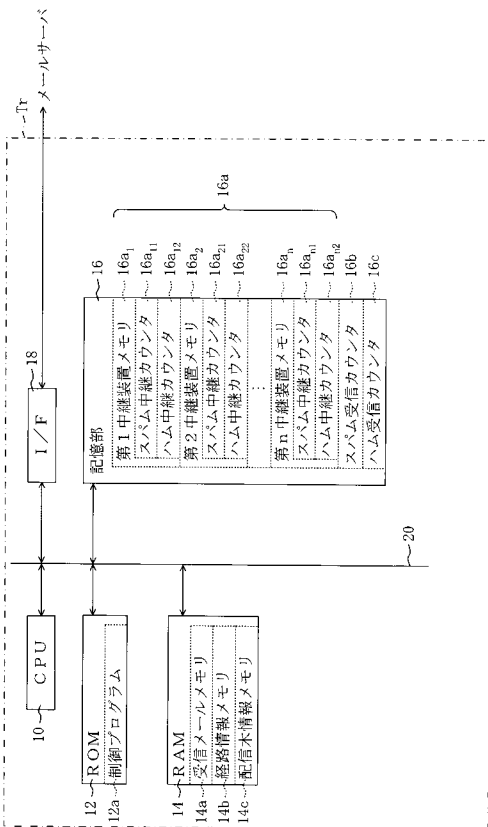


【図2】

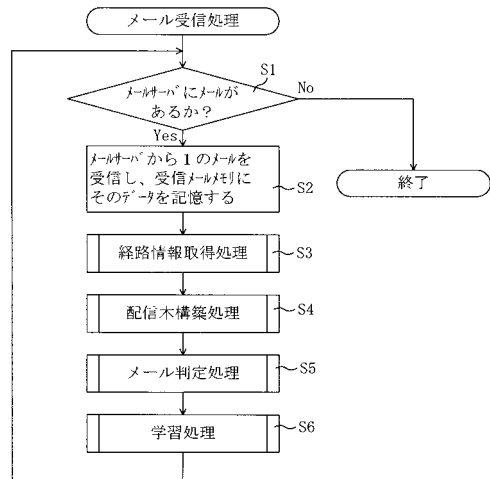
R2 {
 Received: from ceres.xxx.ne.jp (ceres.xxx.ne.jp [211.6.xxx.78])
 by mx1.xxx.or.jp (x.x.x+Sun/x.xW) with ESMTTP id QAA12345:
 Wed, 23 Nov 2005 16:38:08 +0900(JST)

R1 {
 Received: from fmv(p0000-ipadx.xxx.xxx.ne.jp [211.xxx.xxx.156])
 by ceres.xxx.ne.jp (x.x.xx/xxx/) with SMTP id QAA54321:
 Wed, 23 Nov 2005 16:37:08 +0900(JST)

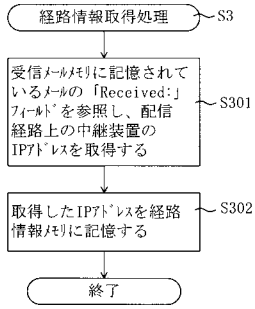
【図3】



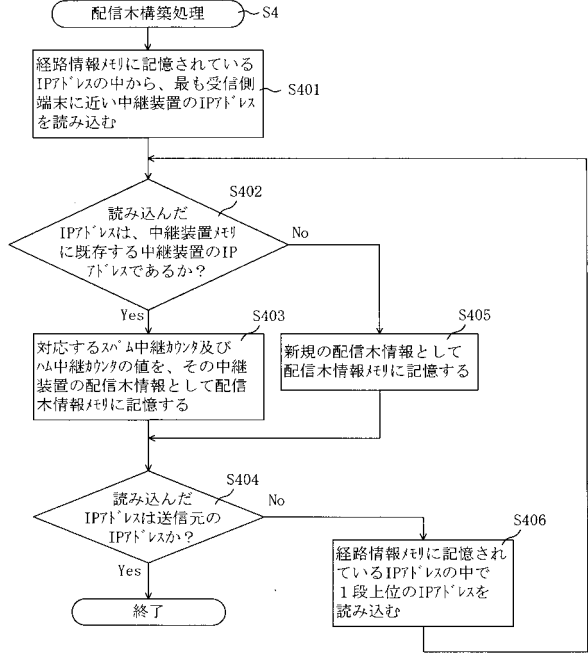
【図4】



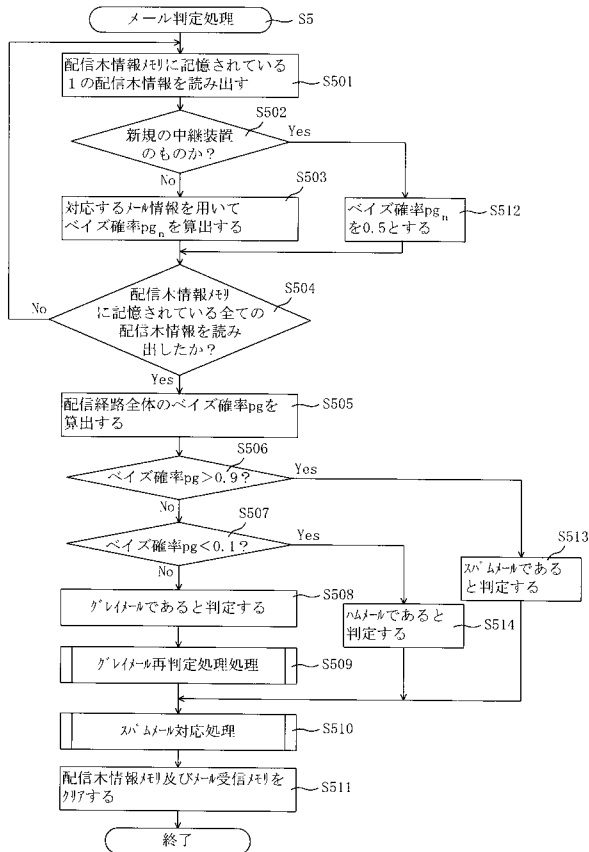
【図5】



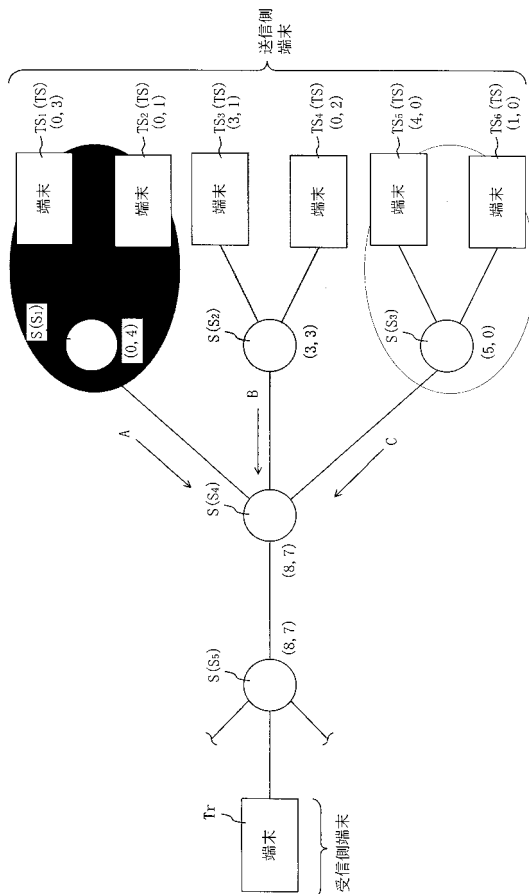
【図6】



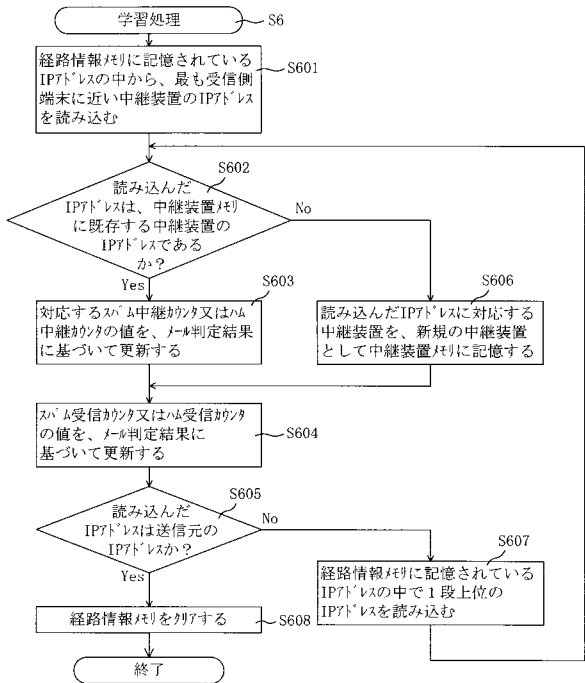
【図7】



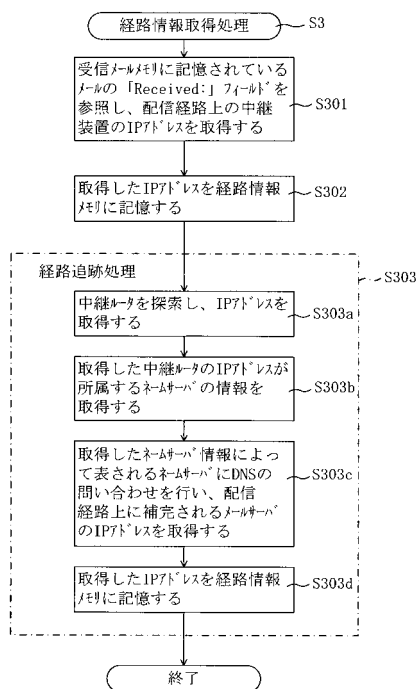
【図8】



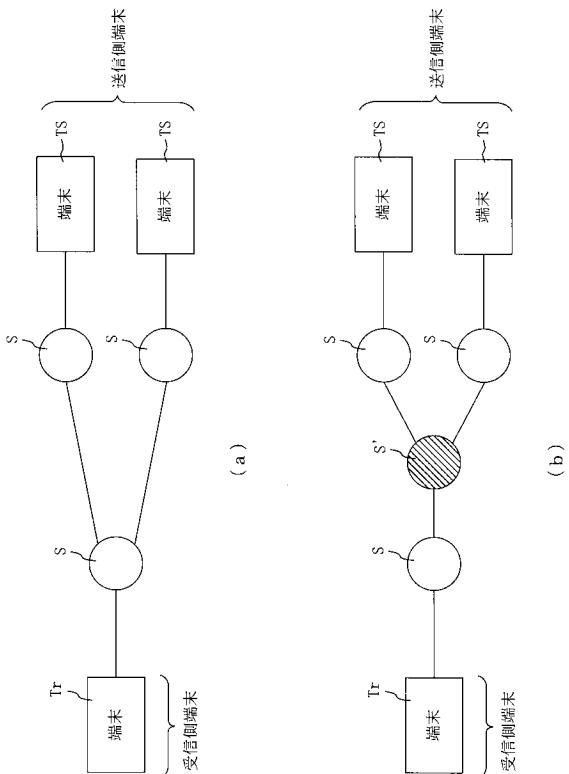
【図9】



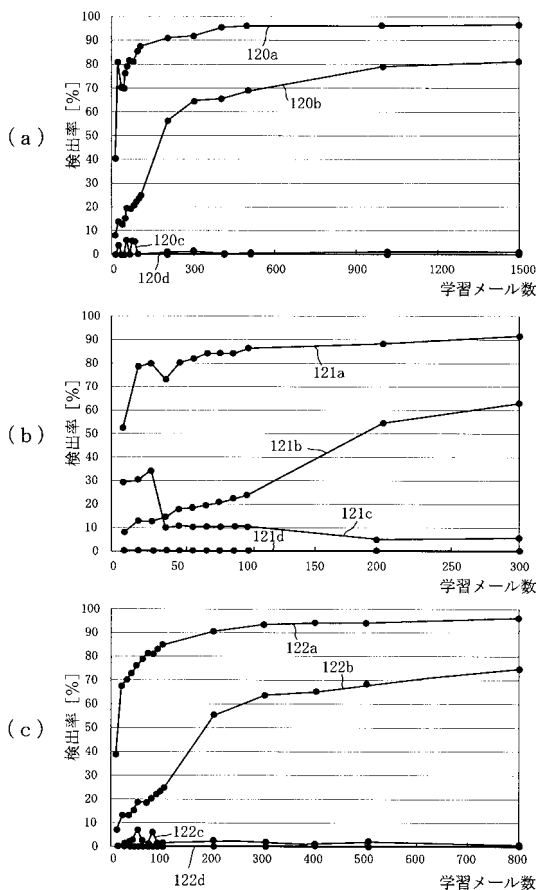
【図10】



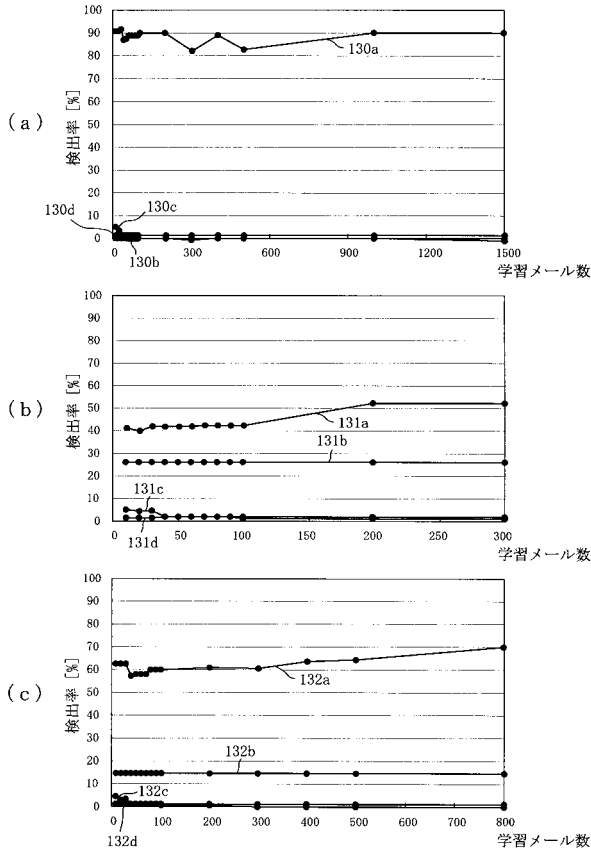
【図11】



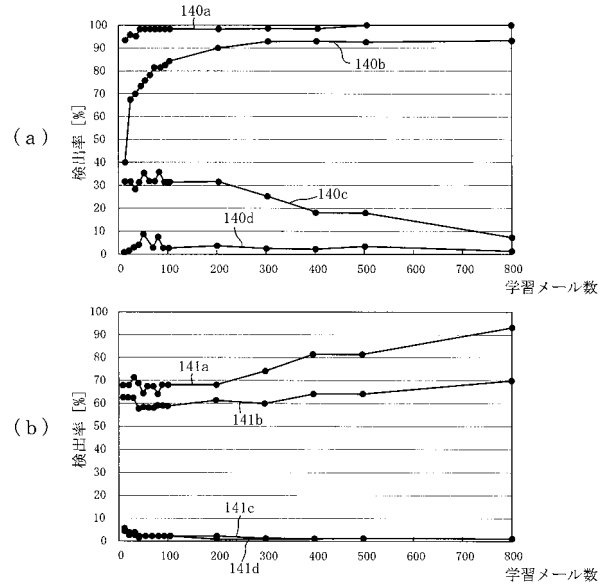
【図12】



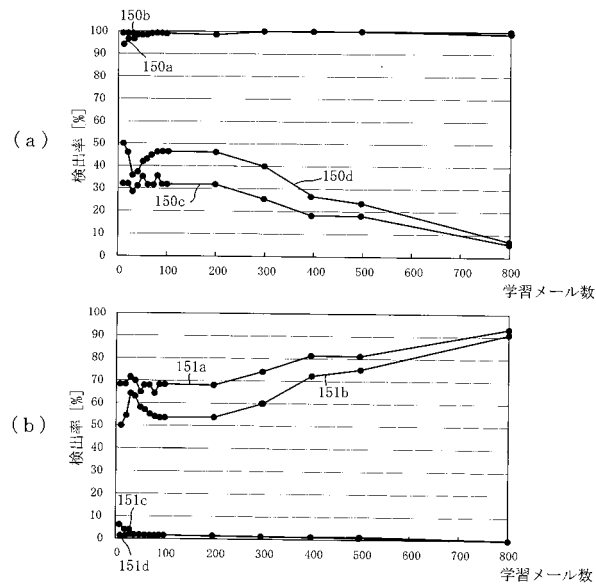
【図13】



【図14】



【図15】



フロントページの続き

(72)発明者 白川 正知

愛知県豊橋市天伯町雲雀ヶ丘1-1

国立大学法人豊橋技術科学大学内

審査官 千本 潤介

(56)参考文献 特開2004-362559(JP,A)

米国特許第06732157(US,B1)

国際公開第2003/071753(WO,A1)

特開2004-015180(JP,A)

中道 理, 基本はフィルタリング, 期待は送信者認証 スпам/フィッシング対策, NIKKEI BYTE 第259号, 日本, 日経BP社, 2004年11月22日, 92~100ページ

井上 孝司, IT業界の定説 嘘か真実か?, COMPUTERWORLD Get Technology Right 第1巻 第5号, 日本, (株)IDGジャパン, 2004年 5月1日, 148~157ページ

(58)調査した分野(Int.Cl., DB名)

G06F 13/00

H04L 12/58